

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:15:37 ; Search time 11.9737 Seconds
(without alignments)
104.437 Million cell updates/sec

Title: US-09-444-281-85

Perfect score: 99

Sequence: 1 ILPWKPWWPWR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96:91526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	100.0	144	JC1222	indolicidin precursor
2	54	54.5	1173	1 VGIHHC	E2 glycoprotein precursor
3	53.5	54.0	299	2 T12505	hypothetical protein
4	53	53.5	327	2 E72851	AcOrf-13 protein - Bom
5	53	53.5	331	2 T41758	AcNPV orf13 - Bom
6	51	51.5	55	2 E90626	ATP synthase F0 chain
7	51	51.5	689	2 AC1927	hypothetical protein
8	51	51.5	1038	2 I38935	bone morphogenetic protein
9	50.5	51.0	970	2 T28234	ORF MSV076 probable
10	50	50.5	83	2 B72392	hypothetical protein
11	50	50.5	337	2 G95922	probable glycosyltransferase
12	50	50.5	498	1 JT0751	ferredoxin-NADP reductase
13	49.5	50.0	296	2 T03562	conserved hypothetical
14	49	49.5	60	2 A56547	sex-peptide precursor
15	49	49.5	425	2 E84631	probable serine carboxylase
16	49	49.5	467	2 E89605	protein F18G5.2 lipase
17	48.5	49.0	111	2 T29295	hypothetical protein
18	48	48.5	55	2 T11105	H+-transporting tw
19	48	48.5	265	2 AH0755	conserved hypothetical
20	48	48.5	400	2 AF2107	hypothetical protein
21	47	47.5	55	1 PWXL8	H+-transporting tw
22	47	47.5	55	2 S68132	H+-transporting tw
23	47	47.5	55	2 S08424	ATP synthase F0 chain
24	47	47.5	55	2 E90618	H+-transporting tw
25	47	47.5	55	2 T11538	H+-transporting tw
26	47	47.5	55	2 T11184	H+-transporting tw
27	47	47.5	55	2 T11291	H+-transporting tw
28	47	47.5	55	2 T09861	H+-transporting tw
29	47	47.5	55	2 T09951	H+-transporting tw

30 47 47.5 55 2 T11768 H+-transporting tw
31 47 47.5 55 2 T11304 H+-transporting tw
32 47 47.5 248 2 S23449 NADH oxidase (H2O2
33 47 47.5 253 2 G70715 hypothetical prote
34 47 47.5 297 2 D87260 integral membrane
35 47 47.5 456 2 T18963 hypothetical prote
36 47 47.5 496 2 A54770 N-acetylglucosamin
37 47 47.5 534 1 S75101 hypothetical prote
38 47 47.5 728 2 T51071 related to trfA pr
39 47 47.5 1112 2 S70522 cyclic nucleotide
40 46.5 47.0 1299 2 AB2244 hypothetical prote
41 46 46.5 54 1 S04619 H+-transporting tw
42 46 46.5 55 2 T11171 H+-transporting tw
43 46 46.5 55 2 T12413 H+-transporting tw
44 46 46.5 55 2 E58892 H+-transporting tw
45 46 46.5 55 2 E90612 ATP synthase F0 ch

ALIGNMENTS

RESULT 1

JC1222

indolicidin precursor - bovine

N/Alternate names: antimicrobial peptide

C/Species: Bos primigenius taurus (cattle)

C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999

C/Accession: JC1222; A42387; S25664

R/del Sal G.; Storici P.; Schneider, C.; Romeo, D.; Zanetti, M.

Biochem. Biophys. Res. Commun. 187, 467-472, 1992

A/Title: cDNA cloning of the neutrophil bactericidal peptide indolicidin.

A/Reference number: JC1222; MUID:92392368; PMID:1520337

A/Accession: JC1222

A/Molecule type: mRNA

A/Residues: 1-144 <SAL>

A/Cross-references: EMBL:X67340; NID:G462; PIDN:CAA47755.1; PID:G463

A/Experimental source: bone marrow

R/Salsed, M.E.; Novotny, M.J.; Morris, W.L.; Tang, Y.Q.; Smith, W.; Cullor, J.S.

J. Biol. Chem. 267, 4292-4295, 1992

A/Title: Indolicidin, a novel bactericidal tridecapeptide amide from neutrophils.

A/Reference number: A42387; MUID:92165771; PMID:1537821

A/Accession: A42387

A/Molecule type: protein

A/Residues: 131-143 <SEL>

A/Experimental source: neutrophils

A/Note: sequence extracted from NCBI backbone (NCBIP:83840)

C/Superfamily: cathelin; cystatin homology

C/Keywords: amidated carboxyl end

F/1-29/Domain: signal sequence #status predicted <SIG>

F/22-129/Domain: cystatin homology <CYS>

F/30-130/Domain: propeptide #status predicted <PRO>

F/131-143/Product: indolicidin #status experimental <MAT>

F/143/Modified site: amidated carboxyl end (Arg) (amide in mature form from following 9

Query Match 100.0%; Score 99; DB 1; Length 144;

Best Local Similarity 100.0%; Pred. No. 7.4e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKPWWPWR 13

Db 131 ILPWKPWWPWR 143

RESULT 2

VGIHHC

E2 glycoprotein precursor - human coronavirus (strain 229E)

N/Alternate names: peplomer glycoprotein; S glycoprotein; spike glycoprotein

C/Species: human coronavirus

A/Note: host Homo sapiens (man)

C/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000

C/Accession: A34166; S05460

R/Raabe, T.; Scheille-Prinz, B.; Siddell, S.G.

J. Gen. Virol. 71, 1065-1073, 1990

A;Title: Nucleotide sequence of the gene encoding the spike glycoprotein of human corona
A;Reference number: A34766; MUID:90264837; PMID:2345367
A;Accession: A34766
A;Molecule type: mRNA
A;Residues: 1-1173 <RA2>
A;Cross-references: EMBL:X15654; NID:958926; PIDN:CAA34723.1; PID:958927
A;Superfamily: coronavirus E2 glycoprotein
C;Keywords: glycoprotein; transmembrane protein
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-1173/Product: E2 glycoprotein #status predicted <MAT>
F;1116-1138/Domain: transmembrane #status predicted <TMN>
F;1132,98,147,171,176,220,243,326,333,440,464,518,538,542,568,581,587,663,671,930,1015,
Query Match 54.5%; Score 54; DB 1; Length 1173;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 KPWPNW 11
DB 1113 KPWPNW 1119
RESULT 3
T12505
Hypothetical protein DKFp434C192.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 23-Jul-1999
C;Accession: T12505
R;Ansoorge, W.; Winkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z17527
A;Accession: T12505
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-299 <ANS>
A;Cross-references: EMBL:AL096753
A;Experimental source: adult testis; clone DKFp434C192
C;Genetics:
A;Note: DKFp434C192.1
Query Match 54.0%; Score 53.5; DB 2; Length 299;
Best Local Similarity 57.1%; Pred. No. 6.4;
Matches 8; Conservative 0; Mismatches 3; Indels 3; Gaps 1;
QY 3 PW---KWPWPWR 13
DB 30 PWYGSAPWPPWR 43
RESULT 4
E72851
AcOf-13 protein - Autographa californica nuclear polyhedrosis virus
C;Species: Autographa californica nuclear polyhedrosis virus, AcMNPV
A;Note: dsDNA virus
C;Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 12-Nov-1999
C;Accession: E72851
R;Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Posse, R.D.
Virology 202, 586-605, 1994
A;Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.
A;Reference number: A72850; MUID:94303173; PMID:8030224
A;Accession: E72851
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-327 <AB>
A;Cross-references: GB:L22858; NID:9510708; PIDN:AAA66643.1; PID:9559082
C;Genetics:
A;Gene: AcOf-13
Query Match 53.5%; Score 53; DB 2; Length 327;
Best Local Similarity 54.5%; Pred. No. 8;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 ILPWKPWPWP 11
DB 1 MLSWLMNWNW 11
RESULT 5
T41758
AcMNPV orf13 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
C;Species: Bombyx mori nuclear polyhedrosis virus, BmNPV
A;Variety: isolate T3
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C;Accession: T41758
R;Gomi, S.; Majima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A;Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
A;Reference number: Z22020; MUID:9281911; PMID:11355780
A;Accession: T41758
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-331 <KAM>
A;Cross-references: EMBL:L33180; NID:93745835; PIDN:AAC63687.1; PID:93745840
A;Experimental source: isolate T3
C;Genetics:
A;Note: Orf_5

Query Match 53.5%; Score 53; DB 2; Length 331;
Best Local Similarity 54.5%; Pred. No. 8.1;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 ILPWKPWPWP 11
DB 1 MLSWLMNWNW 11
RESULT 6
E90626
ATP synthase F0 chain 8 [imported] - Eudromia elegans mitochondrion
C;Species: mitochondrion Eudromia elegans
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 17-May-2002
C;Accession: E90626
R;Haddrath, O.; Baker, A.J.
Proc.R. Soc.Lond. B Biol. Sci. 268, 939-945, 2001
A;Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phyloger
C;Reference number: A99613; MUID:21263106; PMID:11370967
A;Accession: E90626
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-55 <KUR>
A;Cross-references: GB:NC_002772; NID:914141818; PIDN:NP_115277.1; GSPDB:GN00163
C;Genetics:
A;Gene: ATP8
A;Genome: mitochondrion
A;Genetic code: SGI1
C;Superfamily: H+-transporting ATP synthase protein 8
C;Keywords: mitochondrion

Query Match 51.5%; Score 51; DB 2; Length 55;
Best Local Similarity 85.7%; Pred. No. 2.5;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 LPWKWPWP 8
DB 48 LPWSWPWP 54

```

RESULT 7
AC1927
Hypothetical protein all0966 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
R:Accession: AC1927
R:Kaneoko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AC1927
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-689 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA072923.1; PID:gl7130312; GSDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0966

Query Match 51.5%; Score 51; DB 2; Length 689;
Best Local Similarity 55.6%; Pred. No. 29;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Cy 4 WKWPWPWR 12
| | | |
Db 141 WHWGYPWQ 149

RESULT 8
138935
bone morphogenetic protein receptor II precursor - human
N:Alternate names: activin receptor-like kinase type II; bone morphogenetic protein 4 re
N:Contains: protein kinase (SC 2.7.1.37)
C:Species: Homo sapiens (man)
C:Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 24-Sep-1999
A:Accession: I38935; I55438; I37209
R:Kawabata, M.; Chytil, A.; Moses, H.L.
J. Biol. Chem. 270, 5625-5630, 1995
A:Title: Cloning of a novel type II serine/threonine kinase receptor through interaction
A:Reference number: A55947; MUID:95197572; PMID:7890683
A:Accession: I38935
A:Molecule type: mRNA
A:Residues: 1-1038 <KAW>
A:Cross-references: EMBL:U20165; NID:G704361; PIDN:AAC50105.1; PID:G704362
R:Nohno, T.; Ishikawa, T.; Saito, T.; Hosokawa, K.; Noji, S.; Wolsing, D.H.; Rosenbaum,
J. Biol. Chem. 270, 22522-22526, 1995
A:Title: Identification of a human type II receptor for bone morphogenetic protein-4 tha
A:Reference number: I55438; MUID:95403457; PMID:7673243
A:Accession: I55438
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EME
A:Molecule type: mRNA
A:Residues: 1-1038 <NOH>
A:Cross-references: GB:D50516; NID:9807712; PIDN:BA09094.1; PID:G807713
R:Rozenzweig, B.L.; Imamura, T.; Okadome, T.; Cox, G.N.; Yamashita, H.; ten Dijke, P.; H
Proc. Natl. Acad. Sci. U.S.A. 92, 7632-7636, 1995
A:Title: Cloning and characterization of a human type II receptor for bone morphogenetic
A:Reference number: I37209; MUID:95372334; PMID:7644468
A:Accession: I37209
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-827, R', 829-1038 <ROS>
A:Cross-references: EMBL:Z49523; NID:G1009409; PIDN:CAA88759.1; PID:gl009410
C:Genetics:
A:Gene: GDB:BMPPR2; BRX-3; T-ALK; BMPR3; BMPR-II
A:Cross-references: GDB:642243; OMIM:600799
A:Map position: 20pter-20qter
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: ATP; glycoprotein; phosphotransferase; receptor; transmembrane protein
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-1038/Product: bone morphogenetic protein receptor II #status predicted <MAT>

```

```

F:27-150/Domain: extracellular #status predicted <EXT>
F:151-170/Domain: transmembrane #status predicted <TRM>
F:201-508/Domain: protein kinase homology <KIN>
F:209-217/Region: protein kinase ATP-binding motif
F:55,110,126/Binding site: carbonylate (Asn) #status predicted
Query Match 51.5%; Score 51; DB 2; Length 1038;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Cy 3 PWKPWPWP 11
| | | |
Db 8 PWRVPWLPW 16

RESULT 9
T28234
CRF MSV076 probable spheroidin - Melanoplus sanguinipes entomopoxvirus
C:Species: Melanoplus sanguinipes entomopoxvirus
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T28234
R:Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.
J. Virol. 73, 533-552, 1999
A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A:Reference number: Z20484; MUID:99102612; PMID:9847359
A:Accession: T28234
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-970 <AFO>
A:Cross-references: EMBL:AF063866; NID:G4049647; PIDN:AAC97813.1; PID:G4049853
C:Genetics:
A:Note: MSV076

Query Match 51.0%; Score 50.5; DB 2; Length 970;
Best Local Similarity 57.1%; Pred. No. 47;
Matches 8; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

Cy 1 ILPWPWPWP 13
| | | |
Db 838 ILPYYPWPYPYR 851

RESULT 10
B72392
Hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: B72392
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hicke
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.
C.N.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome se
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: B72392
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-83 <ARN>
A:Cross-references: GB:AE001713; GB:AE000512; NID:G4980809; PIDN:AA035403.1; PID:G49808
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0315

Query Match 50.5%; Score 50; DB 2; Length 83;
Best Local Similarity 62.5%; Pred. No. 4.9;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Cy 4 WKWPWPWP 11
| | | |
Db 7 WSWGFWPW 14

RESULT 11

```

G95922
 Probable glycosyltransferase protein Smb21068 [imported] - Sinorhizobium meliloti (strain
 C;Species: Sinorhizobium meliloti
 C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C;Accession: G95922
 R;Pinan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A;Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
 A;Reference number: A95842; MUID:21396508; PMID:11481431
 A;Accession: G95922
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-337 <KUR>
 A;Cross-references: GB:AL591985; PIDN:CA43047.1; PID:G151140532; GSPDB:GN00167
 A;Experimental source: strain 1021, megaplasmid pSymb
 R;Galibert, F.; Pinan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 neault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A;Reference number: A96039; MUID:21368234; PMID:11474104
 A;Contents: annotation
 C;Genetics:
 A;Gene: Smb21068
 A;Genome: plasmid

Query Match 50.5%; Score 50; DB 2; Length 337;
 Best Local Similarity 66.7%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 PWKQPPWPPW 11
 DB 262 PWNTGPPW 270

RESULT 12
 JT0751
 ferredoxin-NADP reductase (EC 1.18.1.2), long form precursor - bovine
 N;Alternate names: adrenodoxin reductase
 C;Species: Bos primigenius taurus (cattle)
 C;Date: 14-Jul-1994 #sequence_revision 18-Oct-1996 #text_change 03-Jun-2002
 C;Accession: JT0751; JT0079; J50390; S03558; P50003; A29604; S52100
 R;Takata, Y.; Sagara, Y.; Kono, A.; Sekimizu, K.; Horiuchi, T.
 Biol. Pharm. Bull. 16, 1200-1206, 1993
 A;Title: Gene structure of bovine adrenodoxin reductase.
 A;Reference number: JT0751; MUID:94177140; PMID:8130767
 A;Accession: JT0751
 A;Molecule type: DNA
 A;Residues: 1-498 <TAK>
 A;Cross-references: GB:D83475; NID:g1199916; PIDN:BA11921.1; PID:g4521308
 A;Experimental source: adrenal cortex
 A;Note: the authors translated the codon GTC for residue 205 as Gly
 R;Sagara, Y.; Takata, Y.; Miyata, T.; Hara, T.; Horiuchi, T.
 J. Biochem. 102, 1333-1336, 1987
 A;Title: Cloning and sequence analysis of adrenodoxin reductase cDNA from bovine adrenal
 A;Reference number: JT0079; MUID:88198050; PMID:3448086
 A;Accession: JT0079
 A;Molecule type: mRNA
 A;Residues: 1-204,211-498 <SAG>
 A;Cross-references: GB:D00211; NID:g217433; PIDN:BA00150.1; PID:g217434
 A;Note: the deduced sequence is partially confirmed by amino acid sequencing of 15 isola
 R;Sagara, Y.
 submitted to DBJ, September 1989
 A;Reference number: J50390
 A;Contents: revision, insertion of residues 205-210
 A;Accession: J50390
 A;Molecule type: mRNA
 A;Residues: 56-498 <SA2>
 R;Hanukoglu, I.; Guttinger, T.
 Eur. J. Biochem. 180, 479-484, 1989
 A;Title: cDNA sequence of adrenodoxin reductase. Identification of NADP-binding sites in

A;Reference number: S03558; MUID:89170752; PMID:2924777
 A;Accession: S03558
 A;Molecule type: mRNA
 A;Residues: 155-204,211-498 <HAN>
 A;Cross-references: EMBL:X13736; NID:g65; PIDN:CAA32002.1; PID:g833776
 A;Note: 405-Ser was also found
 R;Hamamoto, I.; Kurokouchi, K.; Tanaka, S.; Ichikawa, Y.
 Biochim. Biophys. Acta 953, 207-213, 1988
 A;Title: Adrenodoxin-binding peptide of NADPH-adrenodoxin reductase.
 A;Reference number: P50003; MUID:88184054; PMID:3355838
 A;Accession: P50003
 A;Molecule type: Protein
 A;Residues: 33-41,'S',43-62;260-283,'TM';496-498 <HAM>
 A;Note: a cyanogen bromide peptide binds to adrenodoxin
 R;Nonaka, Y.; Murakami, H.; Yabasaki, Y.; Kuramitsu, S.; Kagamiyama, H.; Yamano, T.; Ok
 Biochem. Biophys. Res. Commun. 145, 1239-1247, 1987
 A;Title: Molecular cloning and sequence analysis of full-length cDNA for mRNA of adrenc
 A;Reference number: A29604; MUID:87270696; PMID:3038094
 A;Accession: A29604
 A;Molecule type: mRNA
 A;Residues: 1-76,'R',78-80,'VWLAATTPRSMLL',95-123,'RVYRLT',129-204,211-273,'R',275-322
 A;Cross-references: GB:M17029; NID:g162628; PIDN:AAA30362.1; PID:g162629
 A;Experimental source: adrenal cortex
 R;Waburton, R.J.; Seybert, D.W.
 Biochim. Biophys. Acta 1246, 33-46, 1995
 A;Title: Structural and functional characterization of bovine adrenodoxin reductase by
 A;Reference number: S52100; MUID:95110846; PMID:7811729
 A;Accession: S52100
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: X,34-41,'X',43-48,'X',50-51;304-306,'X',308-309,'X',311-326 <WAR>
 C;Comment: Ferredoxin-NADP+ reductase is localized in the matrix of adrenal cortex mito
 erredoxin-NADP+ reductase, adrenodoxin and two forms of cytochrome P-450.
 C;Genetics:
 A;Introns: 27/1; 59/3; 91/3; 132/3; 170/3; 204/3; 246/3; 275/1; 341/3; 399/1; 456/1
 C;Function:
 A;Description: catalyzes the reversible reduction of NADP+ by reduced ferredoxin or red
 C;Superfamily: human ferredoxin-NADP+ reductase
 C;Keywords: alternative splicing; flavoprotein; mitochondrion; monomer; NADP; oxidoredu
 F11-32/Domain: transit peptide (mitochondrion) #status predicted <SIG>
 F33-498/Product: ferredoxin-NADP+ reductase, long form #status predicted <MAT>
 F33-204,211-498/Product: ferredoxin-NADP+ reductase, short form #status experimental <
 F40-70/Region: beta-alpha-beta FAD nucleotide-binding fold
 F180-190/Region: NADP binding #status predicted
 F728/Binding site: substrate (Lys) #status experimental

Query Match 50.5%; Score 50; DB 1; Length 498;
 Best Local Similarity 66.7%; Pred. No. 28;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 PWKQPPWPPW 11
 DB 3 PRCWRWPPW 11

RESULT 13
 T03562
 conserved hypothetical protein - Rhodobacter capsulatus
 C;Species: Rhodobacter capsulatus
 C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 02-Aug-2002
 C;Accession: T03562
 R;Vlcek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.
 Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997
 A;Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus S910C
 A;Reference number: Z14955; MUID:97404404; PMID:9256491
 A;Accession: T03562
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-296 <VLC>
 A;Cross-references: EMBL:AF010496; NID:g3128256; PIDN:AAC16215.1; PID:g3128363
 C;Genetics:
 A;Map position: 1
 C;Superfamily: hypothetical protein ydeD

```

Matches      7: Conservative      1: Mismatches      2: Indels      0: Gaps      0:

```

Query Match 50.0%; Score 45.3; DB 2; Length 236;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 9; Conservative 3; Mismatches 1; Indels 5; Gaps 2;

Query Macchi	Best Local Similarity	Best Local	pred. No.	pred. No.	score	score
	50.0%	50.0%	20:	20:	49.3;	49.3;

Best Local Similarity: 30.0%, Rec: NO: 20,
Matches 9: Conservative 3: Mismatches 1: Indels 5: Gaps 2:

1 ILPWKWPW-----WP-WRR 13

52 LLPFAWPWLRRDWPLYRR 69

Search completed: May 4, 2004, 15:23:00
Job time : 12.9737 secs

Job time : 12.9737 secs

RESULT 14

A56547
sex-peptide precursor - Drosophila suzukii
N:Alternate names: male accessory gland peptide
C:Species: Drosophila suzukii
C:Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 21-Jul-2000
C:Accession: A56547; B56547
R:Schmidt, T.; Chofrat, Y.; Schneider, M.; Hunziker, P.; Fuyama, Y.; Kubli, E.
Insect Biochem. Mol. Biol. 23, 571-579, 1993
A:Title: Drosophila suzukii contains a peptide homologous to the Drosophila melanogaster
A:Reference number: A56547; MUID:93357785; PMID:835318

Query Match 49.5%; Score 49; DB 2; Length 60;
Best Local Similarity 50.0%; Pred. No. 4.7;
Matches 7; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

Oy 4 WKWPW-----WPWRR 13

20 WEPWYNKOKKPER 33

RESULT 15

E84631
 probable serine carboxypeptidase II [imported] - Arabidopsis thaliana
 C.Species: Arabidopsis thaliana (mouse-ear cress)
 C.Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C.Accession: E84631
 R.Lin.X.: Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 R.Lin.X.: K.S.; Cronin, L.A.; Shen, M.; Varaken, S.E.; Umayam, L.; Tallon, L.;
 K.M.; Koo, H.; Moffat, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 C.; D. Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A.Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

Query Match 49.5%; Score 49; DB 2; Length 425;
Best Local Similarity 70.0%; Pred. No. 32;

A;Map position: 2
C:Superfamily: serine carboxypeptidase

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:08:51 ; Search time 8.21053 Seconds
(without alignments)
82.444 Million cell updates/sec

Title: US-09-444-281-85
Perfect score: 99
Sequence: 1 ILPNKWPWPWRR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	100.0	144	1 INDC_BOVIN	P33046 bos taurus
2	54	54.5	1173	1 VGL2_MYCPU	P15423 human coron
3	53	53.5	327	1 Y013_NPVAC	P41423 autographa
4	51	51.5	1038	1 BMR2_HUMAN	Q13873 homo sapien
5	50	50.5	55	1 ATP8_CORCN	Q9tb16 corythaixoi
6	50	50.5	492	1 ADRO_BOVIN	P08165 bos taurus
7	48	48.5	55	1 ATP8_PELSU	O79674 pelomedusa
8	47.5	48.0	279	1 ELQ1_HUMAN	Q9bw60 homo sapien
9	47.5	48.0	279	1 ELQ1_MOUSE	Q91150 mus musculu
10	47	47.5	55	1 ATP8_GADMO	P15996 gadus morhu
11	47	47.5	55	1 ATP8_ONCMY	P48179 oncorhynch
12	47	47.5	55	1 ATP8_PRODO	Q35416 protopterus
13	47	47.5	55	1 ATP8_SALAL	Q9xn27 salvelinus
14	47	47.5	55	1 ATP8_SALFO	Q9xn35 salvelinus
15	47	47.5	55	1 ATP8_SCYCA	O79405 scyllorhinu
16	47	47.5	55	1 ATP8_SQUAC	Q92250 smalus aca
17	47	47.5	55	1 ATP8_VIRAL	Q9ae31 viroo altil
18	47	47.5	55	1 ATP8_XENLA	P03931 xenopus lae
19	47	47.5	253	1 Y945_MYCTU	P71564 mycobacteri
20	47	47.5	1112	1 CN3B_HUMAN	Q13370 homo sapien
21	46	46.5	54	1 ATP8_CHICK	P44093 gallus gall
22	46	46.5	55	1 ATP8_ANAPL	P50655 anas platyr
23	46	46.5	55	1 ATP8_AYTAM	Q9xkz5 aythya amer
24	46	46.5	55	1 ATP8_CHAPE	Q9tbj9 chaetura pe
25	46	46.5	55	1 ATP8_COLPA	Q9mbd7 columbina p
26	46	46.5	55	1 ATP8_CORCR	Q9tb17 corythaeola
27	46	46.5	55	1 ATP8_COTUA	P50682 coturnix co
28	46	46.5	55	1 ATP8_LATCH	Q31168 latimeria c
29	46	46.5	55	1 ATP8_LOXNO	Q9mdj1 loxigilla n
30	46	46.5	55	1 ATP8_MUSVO	Q9tb18 musophaga v
31	46	46.5	55	1 ATP8_OPINO	Q9tb15 opisthocomu
32	46	46.5	55	1 ATP8_RHEAM	O79396 rheu ameri
33	46	46.5	55	1 ATP8_STRCA	O21401 struthio ca

34	46	46.5	511	1 SYT6_MOUSE	O9ron8 mus musculu
35	45.5	46.0	224	1 Y295_MYCPU	Q98qr6 mycoplasma
36	45	45.5	125	1 YAD3_RHILO	Q98dm8 rhizobium 1
37	45	45.5	337	1 BRB1_RAT	P97583 rattus norv
38	45	45.5	433	1 NHG1_PSEPU	P23262 pseudomonas
39	45	45.5	606	1 ACES_BUNFA	O92035 bungarus fa
40	45	45.5	811	1 PTKS_PSEAE	Q910m3 pseudomonas
41	45	45.5	1108	1 CN3E_RAT	Q63095 rattus norv
42	45	45.5	1154	1 VGL2_IBVD2	P12722 avian infec
43	45	45.5	1162	1 VGL2_IBVK	P1223 avian infec
44	45	45.5	1162	1 VGL2_IBVK	P12650 avian infec
45	45	45.5	1162	1 VGL2_IBVM	P12651 avian infec

ALIGNMENTS

RESULT 1

ID	INDC_BOVIN	STANDARD;	PRT;	144 AA.
AC	P33046;	01-OCT-1993 (Rel. 27, Created)		
DT	01-OCT-1993 (Rel. 27, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Indolicidin precursor.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Bone marrow;			
RX	MEDLINE=92392368; PubMed=1520337;			
RA	del Gal G., Storici P., Schneider C., Romeo D., Zanetti M.;			
RT	"cDNA cloning of the neutrophil bactericidal peptide indolicidin.";			
RL	Biochem. Biophys. Res. Commun. 187:467-472(1992).			
RN	[2]			
RP	SEQUENCE OF 131-143.			
RC	TISSUE=Neutrophils;			
RX	MEDLINE=92165771; PubMed=1537821;			
RA	Selsted M.E., Novotny M.J., Morris W.L., Tang Y.-Q., Smith W.;			
RT	Cullor J.S.;			
RL	"Indolicidin, a novel bactericidal tridecapeptide amide from neutrophils.";			
CC	J. Biol. Chem. 267:4292-4295(1992).			
CC	-I- FUNCTION: Potent microbicidal activity, active against Staphylococcus aureus and Escherichia coli.			
CC	-I- TISSUE SPECIFICITY: Large granules of neutrophils.			
CC	-I- PTM: Elastase might be responsible for its maturation.			
CC	-I- SIMILARITY: Belongs to the cathelicidin family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; X67340; CAA47755.1; -.			
DR	PIR; JCI222; JCI222.			
DR	PDB; 1G89; 17-JAN-01.			
DR	PDB; 1G8C; 17-JAN-01.			
DR	PDB; 1HRI; 21-DEC-02.			
DR	InterPro; IPR001894; Cathelicidin.			
DR	ProDom; PD001838; Cathelicidin; 1.			
DR	ProSITE; PS00946; CATHELICIDINS_1; 1.			
DR	ProSITE; PS00947; CATHELICIDINS_2; 1.			
KW	Antibiotic; Amidation; Signal; Pyrrolidone carboxylic acid;			
KW	3D-structure.			
FT	SIGNAL 1 29 POTENTIAL.			

```

FT PROPP 30 130
FT PEPTIDE 131 143
FT MOD_RES 30 30
FT INDOLICIDIN.
FT PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY)
FT DISULFID 85 96
FT BY SIMILARITY.
FT DISULFID 107 124
FT BY SIMILARITY.
FT MOD_RES 143 143
FT AMIDATION (G-144 PROVIDE AMIDE GROUP).
SQ SEQUENCE 144 AA; 16479 MW; E3B1CBCE55C09911 CRC64;

Query Match 100.0%; Score 99; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILPWKPWPWRR 13
Db 131 ILPWKPWPWRR 143

RESULT 2
VGL2_CVH22
ID VGL2_CVH22 STANDARD; PRT; 1173 AA.
AC P15423; P89342; P89343; P89344; Q66174; Q990M1; Q990M2; Q990M3;
AC Q990M4;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein).
GN S.
OS Human coronavirus (strain 229E) (HCoV-229E).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OC NCB_TaxID=11137;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=90264837; PubMed=2345367;
RA Raabe T., Schelle-Prinz B., Siddell S.G.;
RT "Nucleotide sequence of the gene encoding the spike glycoprotein of
human coronavirus HCV 229E."
RT J. Gen. Virol. 71:1065-1073(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21262210; PubMed=11369870;
RA Thiel V., Herold J., Schelle B., Siddell S.G.;
RT "Infectious RNA transcribed in vitro from a cDNA copy of the human
coronavirus genome cloned in vaccinia virus."
RT J. Gen. Virol. 82:1273-1281(2001).
RN [3]
RP SEQUENCE OF 98-1113 FROM N.A., AND VARIANTS.
RC STRAIN=Isolate RW Stock, Isolate P100E, Isolate P11A, and
RC Isolate P11B;
RA Bonavia A., Holmes K.V.;
RT "Viral and cellular changes in a human cell line persistently infected
with human coronavirus HCoV-229E."
RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 98-1113 FROM N.A., AND VARIANTS.
RC STRAIN=Isolate ATCC VR-74, Isolate A162, and Isolate LRI 281;
RX MEDLINE=99086140; PubMed=9870593;
RA Hays J.P., Myint S.H.;
RT "PCR sequencing of the spike genes of geographically and
chronologically distinct human coronaviruses 229E."
RT J. Virol. Methods 75:179-193(1998).
RN [5]
RP SEQUENCE OF 1159-1173 FROM N.A.
RX MEDLINE=89366667; PubMed=2701946;
RA Raabe T., Siddell S.;
RT "Nucleotide sequence of the human coronavirus HCV 229E mRNA 4 and mRNA
5 unique regions."
RN [6]
RL Nucleic Acids Res. 17:6387-6387(1989).
RP INTERACTION WITH ANPEP.
RX MEDLINE=22440020; PubMed=12551991;
RA Bonavia A., Zelus B.D., Wentworth D.E., Talbot P.J., Holmes K.V.;

```

```

RT "Identification of a receptor-binding domain of the spike glycoprotein
of human coronavirus HCoV-229E."
J. Virol. 77:2530-2538(2003).
RN [7]
RP INTERACTION WITH ANPEP.
RX MEDLINE=22521439; PubMed=12634402;
RA Breslin J.J., Mork I., Smith M.K., Vogel L.K., Hemmila E.M.,
RA Bonavia A., Talbot P.J., Sjoestrom H., Noren O., Holmes K.V.;
RT "Human coronavirus 229E: receptor binding domain and neutralization by
soluble receptor at 37 degrees C."
J. Virol. 77:4435-4438(2003).
RN [8]
RP REVIEW.
RX MEDLINE=21109095; PubMed=11162792;
RA Gallagher T.M., Buchmeier M.J.;
RT "Coronavirus spike proteins in viral entry and pathogenesis."
Virology 279:371-374(2001).
CC -1- FUNCTION: Structural protein that makes spikes at the surface of
the virus. Determines enteropathogenicity and virulence of the
virus. Initiates infection by specifically recognizing and binding
the human aminopeptidase ANPEP receptor. Its association with
ANPEP may lead to its conformational change that triggers fusion
between viral and host cellular membrane.
CC -1- SUBUNIT: Homotrimer. During virus morphogenesis, it is found in a
complex with M and HE proteins (By similarity). Interacts with
ANPEP.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DOMAIN: The spike S1 domain displays the specificity for the host
receptor.
CC -1- DOMAIN: The leucine zipper-like heptad repeats may mediate the
fusion of viral and cellular membranes.
CC -1- POLYMORPHISM: The strong variation between the different
strains may affect the virulence of the virus.
CC -1- MISCELLANEOUS: In contrast to serogroup 2, E2 glycoprotein protein
from serogroup 1 is not cleaved.
CC -1- SIMILARITY: Contains 1 spike S1 domain.
CC -1- SIMILARITY: Contains 1 spike S2 domain.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
EMBL; X16816; CAA34723.1; -
EMBL; AF304460; AAG48592.1; -
EMBL; AF344186; AAK32188.1; -
EMBL; AF344187; AAK32189.1; -
EMBL; AF344188; AAK32190.1; -
EMBL; AF344189; AAK32191.1; -
EMBL; Y09923; CAA71056.1; -
EMBL; Y10051; CAA71146.1; -
EMBL; Y10052; CAA71147.1; -
EMBL; X15654; CAA33680.1; -
PIR; A34766; VGIHHC.
DR InterPro; IPR002551; Corona_S1.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01600; Corona_S1; 1.
DR Pfam; PF01601; Corona_S2; 1.
DR Virulence; Glycoprotein; Envelope protein; Transmembrane; Signal;
KW Coiled coil.
FT SIGNAL 1 15
FT CHAIN 16 1173
FT DOMAIN 16 1115
FT POTENTIAL
FT TRANSMEM 1116 1135
FT DOMAIN 1136 1173
FT DOMAIN 32 536
FT DOMAIN 417 547
FT DOMAIN 537 1171
FT DOMAIN 1054 1103
FT DOMAIN 1067 1102
FT E2 GLYCOPROTEIN.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT SPIKE S1.
FT INTERACTION WITH ANPEP.
FT SPIKE S2.
FT COILED COIL (POTENTIAL).
FT LEUCINE ZIPPER-LIKE HEPTAD REPEATS.

```


GN BMPR2 OR PPH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1] SEQUENCE FROM N.A.
RP TISSUE=Substantia nigra;
RC MEDLINE=95372334; PubMed=7644468;
RX Rosenzweig B.L., Imamura T., Okadome T., Cox G.N., Yamashita H.,
RA ten Dijke P., Heldin C., Miyazono K.;
RT "Cloning and characterization of a human type II receptor for bone
RT morphogenetic proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:7632-7636(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin fibroblast;
RX MEDLINE=95403457; PubMed=7673243;
RA Nohno T., Iehikawa T., Saito T., Hosokawa K., Noji S., Wosing D.H.,
RA Rosenbaum J.S.;
RT "Identification of a human type II receptor for bone morphogenetic
RT protein-4 that forms differential heteromeric complexes with bone
RT morphogenetic protein type I receptors.";
RL J. Biol. Chem. 270:22522-22526(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95197572; PubMed=7890683;
RA Kawabata M., Chytil A., Moses H.L.;
RT "Cloning of a novel type II serine/threonine kinase receptor through
RT interaction with the type I transforming growth factor-beta
RT receptor.";
RL J. Biol. Chem. 270:5625-5630(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin.
RX MEDLINE=22368257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP VARIANTS PPH1 GLN-491 AND TRP-491.
RX MEDLINE=20395844; PubMed=1093931;
RA Deng Z., Morse J.H., Slager S.L., Cuervo N., Moore K.J., Venetos G.,
RA Kalachikov S., Cayanis E., Fischer S.G., Barst R.J., Hodge S.E.,
RA Knowles J.A.;
RT "Familial primary pulmonary hypertension (gene PPH1) is caused by
RT mutations in the bone morphogenetic protein receptor-II gene.";
RL Am. J. Hum. Genet. 67:737-744(2000).
RN [6]
RP VARIANTS PPH1 TYR-60; TYR-117 AND ARG-483.
RX MEDLINE=20473811; PubMed=11015450;
RA Thomson J.R., Machado R.D., Paucullo M.W., Morgan N.V., Humbert M.,
RA Elliott G.C., Ward K., Yacoub M., Mikhail G., Rogers P., Newman J.H.,
RA Wheeler L., Higenbottam T., Gibbs J.S.R., Egan J., Crozier A.,
RA Peacock A., Allcock R., Corris P., Loyd J.E., Trembath R.C.,
RA Nichols W.C.;

"Sporadic primary pulmonary hypertension is associated with germline
mutations of the gene encoding BMPR-II, a receptor member of the
TGF-beta family.";
RL J. Med. Genet. 37:741-745(2000).
RN [7]
RP VARIANTS PPH1 TRP-118; TYR-347 AND GLY-485.
RX MEDLINE=20428187; PubMed=10973254;
RA Lane K.B., Machado R.D., Paucullo M.W., Thomson J.R., Aldred M.,
RA Phillips J.A. III, Loyd J.E., Nichols W.C., Trembath R.C., Gaddipati R.,
RA Brannon C.A., Conneally P.M., Foroud T., Fretwell N., Gaddipati R.,
RA Koller D., Loyd E.J., Morgan N.V., Newman J.H., Prince M.A.,
RA Villalón Guesell C., Wheeler L.;
RT "Heterozygous germline mutations in BMPR2, encoding a TGF-beta
RT receptor, cause familial primary pulmonary hypertension.";
RL Nat. Genet. 26:81-84(2000).
RN [8]
RP VARIANTS PPH1 ARG-123; SER-123; ARG-420 AND THR-512, VARIANT ASP-224,
RX MEDLINE=21063176; PubMed=11115378;
RA Machado R.D., Paucullo M.W., Thomson J.R., Lane K.B., Morgan N.V.,
RA Wheeler L., Phillips J.A. III, Newman J.H., Williams D., Galie N.,
RA Manes A., McNeil K., Yacoub M., Mikhail G., Rogers P., Corris P.,
RA Humbert M., Donnai D., Martensson G., Trembath R.C., Loyd J.E.,
RA Trembath R.C., Nichols W.C.;
RT "BMPR2 haploinsufficiency as the inherited molecular mechanism for
RT primary pulmonary hypertension.";
RL Am. J. Hum. Genet. 68:92-102(2001).
CC -1- FUNCTION: Binds to BMP-7, BMP-2 and, less efficiently, BMP-4.
CC Binding is weak but enhanced by the presence of type I receptors
CC for BMPs.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SUBUNIT: Heterodimer with type-I receptors.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Highly expressed in heart and liver.
CC -1- DISEASE: Defects in BMPR2 are the cause of primary pulmonary
CC hypertension (PPH1) [MIM:178600]; a rare autosomal dominant
CC disorder characterized by plexiform lesions of proliferating
CC endothelial cells in pulmonary arterioles. The lesions lead to
CC elevated pulmonary arterial pressure, right ventricular failure,
CC and death. The disease can occur from infancy throughout life and
CC it has a mean age at onset of 36 years. Penetrance is reduced.
CC Although familial PPH1 is rare, cases secondary to known
CC etiologies are more common and include those associated with the
CC appetite-suppressant drugs.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. TGFB
CC receptor subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; Z48923; CAA88759.1; -;
EMBL; D50516; BAA03094.1; -;
EMBL; U20165; AAC50105.1; -;
EMBL; BC052985; AAB52985.1; -;
PIR; I38935; I38935.
Gene; HGNC:1078; BMPR2.
MIM; 600799; -;
MIM; 178600; -;
GO; GO:0005887; C:integral to plasma membrane; TAS.
GO; GO:0005515; P:protein binding; TAS.
GO; GO:0007178; P:transmembrane receptor protein serine/threonine kinase; TAS.
InterPro; IPR000472; Activin_receptor.
InterPro; IPR000719; Prot_kinase.
Pfam; PF01064; Ser_thr_kinase.
Pfam; PF00069; pkinase; 1.
ProDom; PD000001; Prot_kinase; 1.
ProSITE; PS00107; PROTEIN_KINASE_ATP; 1.

[4] SEQUENCE FROM N.A.
 RP TISSUE=Adrenal cortex;
 RX MEDLINE=89170752; PubMed=2924777;
 RA Hanukoglu I., Gutfinger T.;
 RT "cDNA sequence of adrenodoxin reductase. Identification of NADP-
 binding sites in oxidoreductases.";
 RL Eur. J. Biochem. 180:479-484(1989).
 RN [5]
 RP SEQUENCE OF N-TERMINUS, AND PARTIAL SEQUENCE.
 RC TISSUE=Adrenal cortex;
 RX MEDLINE=88082777; PubMed=3691502;
 RA Hanukoglu I., Gutfinger T., Hanlu M., Shively J.E.;
 RT "Isolation of a cDNA for adrenodoxin reductase (ferredoxin-NADP+
 reductase). Implications for mitochondrial cytochrome P-450 systems.";
 RL Eur. J. Biochem. 169:449-455(1987).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 33-492.
 RC TISSUE=Adrenal gland;
 RX MEDLINE=99299392; PubMed=10369776;
 RA Ziegler G.A., Vonrhein C., Hanukoglu I., Schulz G.E.;
 RT "The structure of adrenodoxin reductase of mitochondrial P450 systems:
 electron transfer for steroid biosynthesis.";
 RL J. Mol. Biol. 289:981-990(1999).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
 RX MEDLINE=20455764; PubMed=1098235;
 RA Ziegler G.A., Schulz G.E.;
 RT "Crystal structures of adrenodoxin reductase in complex with NADP+ and
 NADPH suggesting a mechanism for the electron transfer of an enzyme
 family.";
 RL Biochemistry 39:10986-10995(2000).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH ADRENODOXIN.
 RX MEDLINE=21264735; PubMed=11053423;
 RA Mueller J.J., Lapko A., Bourenkov G., Ruckpaul K., Heinemann U.;
 RT "Adrenodoxin reductase-adrenodoxin complex structure suggests electron
 transfer path in steroid biosynthesis.";
 RL J. Biol. Chem. 276:2786-2789(2001).
 CC -1- FUNCTION: Serves as the first electron transfer protein in all the
 mitochondrial P450 systems. Including cholesterol side chain
 cleavage in all steroidogenic tissues, steroid 11-beta
 hydroxylation in the adrenal cortex, 25-OH-vitamin D3-24
 hydroxylation in the kidney, and steroid C-27 hydroxylation in the
 liver.
 CC -1- CATALYTIC ACTIVITY: Reduced adrenodoxin + NADP(+) = oxidized
 adrenodoxin + NADPH.
 CC -1- COFACTOR: FAD.
 CC -1- PATHWAY: Cholesterol side-chain-cleavage system.
 CC -1- SUBUNIT: Monomer.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Short;
 CC IsoId=P08165-1; Sequence=Displayed;
 CC Name=Long;
 CC IsoId=P08165-2; Sequence=VSP_003415;
 CC Note=Represents 10-20% of all adrenodoxin reductase mRNAs and
 seems to be inactive;
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D83475; BAAL1921.1; -;
 DR EMBL; D83472; BAAL1921.1; JOINED.
 DR EMBL; D83473; BAAL1921.1; JOINED.
 DR EMBL; D83474; BAAL1921.1; JOINED.
 DR EMBL; M17029; AAA30362.1; -;

DR EMBL; D00211; BAA00150.1; -;
 DR EMBL; X13736; CAA32002.1; -;
 DR PIR; JT0751; JT0751.
 DR PDB; 1CJC; 12-APR-99.
 DR PDB; 1E1L; 24-SEP-00.
 DR PDB; 1E1L; 24-SEP-00.
 DR PDB; 1E1K; 24-SEP-00.
 DR PDB; 1E1M; 24-SEP-00.
 DR PDB; 1E1N; 24-SEP-00.
 DR PDB; 1E6E; 01-AUG-03.
 DR InterPro; IPR000759; Adrndx_reductase.
 DR PRINTS; PR00419; ADXRDTASE.
 DR Electron transport; Cholesterol metabolism; Oxidoreductase;
 KW Mitochondrion; FAD; Flavoprotein; NADP; Transit peptide;
 KW Alternative splicing; 3D-structure.
 FT TRANSIT 1 32 MITOCHONDRION.
 FT CHAIN 33 492 NADPH:ADRENODOXIN OXIDOREDUCTASE.
 FT VARSPLIC 204 204 E -> EVLLCQ (in isoform Long).
 FT CONFLICT 77 77 /FTid=VSP_003415.
 FT CONFLICT 81 94 G -> R (IN REF. 3).
 FT FGVAPEHPEVKNI -> VWLALTTPSRMLL (IN REF. 3).
 FT QDAYH -> RYVRLT (IN REF. 3).
 FT K -> R (IN REF. 3).
 FT PS -> RL (IN REF. 3).
 FT RAAGIRLAVTR -> ARRSAWQSP (IN REF. 3).
 FT TRAVPTGDVEDL -> HPGSAHWGCGGP (IN REF. 3).

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF151846; AAD34083.1; ALT_FRAME.
DR EMBL; AK001653; BAA91813.1; -.
DR EMBL; BC000618; AAH00618.1; -.
DR Genew; HGNC:14418; ELOVL1_SUR4.
DR InterPro; IPR002076; GNS1_SUR4.
DR Pfam; PF01151; ELO; 1.
DR PROSITE; PS01188; ELO; 1.
KW Fatty acid biosynthesis; Transmembrane; Endoplasmic reticulum.
FT TRANSMEM 23 43 POTENTIAL.
FT TRANSMEM 61 81 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 201 221 POTENTIAL.
FT TRANSMEM 231 251 POTENTIAL.
FT SITE 275 277 POTENTIAL.
FT CONFLICT 68 68 S -> P (IN REF. 2).
SQ SEQUENCE 279 AA; 32663 MW; B168EE4C7EAF92A6 CRC64;

Query Match 48.0%; Score 47.5; DB 1; Length 279;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 ILPWKPFW 9

Db 147 VLPWSW-WW 154

RESULT 9

ID ELO1 MOUSE STANDARD; PRT; 279 AA.
AC Q9JL5; Q9JL52.
DT 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Elongation of very long chain fatty acids protein 1.
GN ELOVL1 OR SSC1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RX MEDLINE=20253178; PubMed=10791983;
RA Tvrdek P., Westerberg R., Silve S., Asadi A., Jakobsson A., Cannon B.,
RA Lolson G., Jacobsson A.;
RA "Role of a new mammalian gene family in the biosynthesis of very long
RA chain fatty acids and sphingolipids.";
RA J. Cell Biol. 149:707-718 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]

RP SEQUENCE OF 78-279 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Komno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G.,
RA Blake J., Bojelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Iyoy-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).

CC -1- FUNCTION: Could be implicated in tissue-specific synthesis of very
CC long chain fatty acids and sphingolipids. May catalyze one or both
CC of the reduction reaction in fatty acid elongation, i.e.,
CC conversion of beta-ketoacyl CoA to beta-hydroxyacyl CoA or
CC reduction of trans-2-enoyl CoA to the saturated acyl CoA
CC derivative.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum (Potential).

CC -1- TISSUE SPECIFICITY: Expressed in a broad variety of tissues.
CC Highly expressed in stomach, lung, kidney, skin and intestine.
CC Moderately expressed in white adipose tissue, liver, spleen, and in
CC brain, brown adipose tissue, heart and muscle. Weakly expressed in
CC testis.

CC -1- SIMILARITY: Belongs to the ELO family.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL; AF170907; AAF72572.1; -.

CC EMBL; BC006735; AAH06735.1; -.

CC EMBL; AK003743; BAB22975.1; -.

CC MGD; MGI:1858959; Elovil.

CC InterPro; IPR002076; GNS1_SUR4.

CC Pfam; PF01151; ELO; 1.

CC PROSITE; PS01188; ELO; 1.

DR Fatty acid biosynthesis; Transmembrane; Endoplasmic reticulum.

DR TRANSMEM 23 43 POTENTIAL.

DR TRANSMEM 61 81 POTENTIAL.

DR TRANSMEM 176 196 POTENTIAL.

DR TRANSMEM 203 223 POTENTIAL.

DR TRANSMEM 231 251 POTENTIAL.

DR SITE 275 277 POTENTIAL.

DR CONFLICT 78 79 YE -> MR (IN REF. 3).

SQ SEQUENCE 279 AA; 32678 MW; CA5A1CF55FDB2F76 CRC64;

Query Match 48.0%; Score 47.5; DB 1; Length 279;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

```

QY      1 ILPNKWPWW 9
Db      147 VLPWSW-WW 154

RESULT 10
ATP8_GADMO
ID      ATP8_GADMO      STANDARD;      PRT;      55 AA.
AC      P15996;
DT      01-APR-1990 (Rel. 14, Created)
DT      01-APR-1990 (Rel. 14, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN      MTATP8 OR ATP8.
OS      Gadus morhua (Atlantic cod).
OG      Mitochondrion.
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC      Agnathomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX      NCBI_TaxID=8049;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Norwegian coastal 1; TISSUE=Liver;
RX      MEDLINE=90174958; PubMed=2308841;
RA      Johansen S., Guddal P.H., Johansen T.;
RT      "Organization of the mitochondrial genome of Atlantic cod, Gadus
RT      morhua.";
RL      Nucleic Acids Res. 19:411-419(1990).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Norwegian coastal 1;
RX      MEDLINE=96414925; PubMed=8817926;
RA      Johansen S., Bakke I.;
RT      "The complete mitochondrial DNA sequence of Atlantic cod (Gadus
RT      morhua): relevance to taxonomic studies among codfishes.";
RL      Mol. Mar. Biol. Biotechnol. 5:203-214(1996).
CC      -!- FUNCTION: This is one of the chains of the nonenzymatic component
CC      (CF(0) subunit) of the mitochondrial ATPase complex.
CC      -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC      H(+) (Out).
CC      -!- SUBCELLULAR LOCATION: Membrane-bound.
CC      -!- SIMILARITY: Belongs to the ATPase protein 8 family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X17659; CAA35655.1; -
DR      EMBL; X99772; CAA68110.1; -
DR      PIR; S08424; S08424.
DR      InterPro; IPR001421; ATPase8_mit.
DR      Pfam; PF00895; ATP-synt_8; 1.
KW      Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT      TRANSMEM 4 24 POTENTIAL.
SQ      SEQUENCE 55 AA; 6481 MW; B85C81E63DB48B15 CRC64;

Query Match 47.5%; Score 47; DB 1; Length 55;
Best Local Similarity 83.3%; Pred. No. 4.9;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 PWKWPW 8
Db      49 PWNWEP 54

RESULT 11
ATP8_ONCMY
ID      ATP8_ONCMY      STANDARD;      PRT;      55 AA.

P48179;
01-FEB-1996 (Rel. 33, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
MTATP8 OR ATP8.
Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri), and
Salmo salar (Atlantic salmon).
Mitochondrion.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Paracanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
NCBI_TaxID=8022, 8030;
[1]
SEQUENCE FROM N.A.
SPECIES=O.mykiss; TISSUE=Liver;
MEDLINE=96139077; PubMed=8587139;
Zardoya R., Garrido-Perterra A., Bautista J.M.;
"The complete nucleotide sequence of the mitochondrial DNA genome of
the rainbow trout, Oncorhynchus mykiss.";
J. Mol. Evol. 41:942-951(1995).
[2]
SEQUENCE FROM N.A.
SPECIES=S.salar; TISSUE=Liver;
MEDLINE=20018174; PubMed=10548724;
Hurst C.D., Bartlett S.E., Davidson W.S., Bruce I.J.;
"The complete mitochondrial DNA sequence of the Atlantic salmon, Salmo
salar.";
Gene 239:237-242(1999).
[3]
SEQUENCE FROM N.A.
SPECIES=S.salar;
Arnaason U., Johnsson E., Rasmussen A.S.;
"The complete mitochondrial genome sequence of a teleost, Salmo salar,
and comparisons with other salmoniformes.";
Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: This is one of the chains of the nonenzymatic component
(CF(0) subunit) of the mitochondrial ATPase complex.
-!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
H(+) (Out).
-!- SUBCELLULAR LOCATION: Membrane-bound.
-!- SIMILARITY: Belongs to the ATPase protein 8 family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR      EMBL; L29771; AAB03351.1; -
DR      EMBL; U12143; AAD04737.1; -
DR      EMBL; AF133701; AAF61382.1; -
DR      PIR; T09861; T09861.
DR      PIR; T09951; T09951.
DR      InterPro; IPR001421; ATPase8_mit.
DR      Pfam; PF00895; ATP-synt_8; 1.
KW      Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT      TRANSMEM 4 24 POTENTIAL.
SQ      SEQUENCE 55 AA; 6413 MW; D02920C3E3346925F CRC64;

Query Match 47.5%; Score 47; DB 1; Length 55;
Best Local Similarity 83.3%; Pred. No. 4.9;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 PWKWPW 8
Db      49 PWNWEP 54

RESULT 12
ATP8_PRODO

```

```

ID  ATP8_PRODO      STANDARD;      PRT;      55 AA.
AC  Q35416;
DT  15-JUL-1999 (Rel. 38, Created)
DT  15-JUL-1999 (Rel. 38, Last sequence update)
DT  15-JUL-1999 (Rel. 43, Last annotation update)
DE  ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN  MTATP8 OR ATP8
OS  Protopterus dolloi (Slender lungfish).
OC  Mitochondrion.
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Dipnoi; Lepidosireniformes; Protopterygidae; Protopterus.
OX  NCBI_TaxID=27779;
RN  [1]_TaxID=27779;
RP  SEQUENCE FROM N.A.
RC  TISSUE=Egg;
RX  MEDLINE=96271539; PubMed=8846902;
RA  Zardoya R., Meyer A.;
RT  "The complete nucleotide sequence of the mitochondrial genome of the
RT  lungfish (Protopterus dolloi) supports its phylogenetic position as a
RT  close relative of land vertebrates.";
RL  Genetics 142:1249-1263 (1996).
CC  -!- FUNCTION: This is one of the chains of the nonenzymatic component
CC  -!- (CF(0) subunit) of the mitochondrial ATPase complex.
CC  -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC  H(+) (Out).
CC  -!- SUBCELLULAR LOCATION: Membrane-bound.
CC  -!- SIMILARITY: Belongs to the ATPase protein 8 family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  EMBL; L42813; AAC38025.1; -.
DR  PIR; S68132; S68132.
DR  InterPro; IPR001421; ATPase8_mit.
DR  Pfam; PF00895; ATP-synt_8; 1.
KW  Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
SQ  SEQUENCE 55 AA; 6523 MW; 95343043B5B2DC53 CRC64;

Query Match 47.5%; Score 47; DB 1; Length 55;
Best Local Similarity 83.3%; Pred. No. 4.9;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  3 PWKWPW 8
DB  49 PWNWPW 54

RESULT 13
ATP8_SALAL
ID  ATP8_SALAL      STANDARD;      PRT;      55 AA.
AC  Q9XN27;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  10-OCT-2003 (Rel. 42, Last annotation update)
DE  ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN  MTATP8 OR ATP8 OR ATPASE8
OS  Salvelinus alpinus (Arctic char).
OC  Mitochondrion.
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC  Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX  NCBI_TaxID=8036;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Doiron S., Blier P.U., Bernatchez L.;
RT  "A comparative analysis of complete sequence of mitochondrial genome
RT  between brook char (Salvelinus fontinalis) and arctic char (S.
RT  alpinus).";
RL  Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC  -!- FUNCTION: This is one of the chains of the nonenzymatic component
CC  -!- (CF(0) subunit) of the mitochondrial ATPase complex.
CC  -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC  H(+) (Out).
CC  -!- SUBCELLULAR LOCATION: Membrane-bound.
CC  -!- SIMILARITY: Belongs to the ATPase protein 8 family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  EMBL; AF154850; AAD41376.1; -.
DR  InterPro; IPR001421; ATPase8_mit.
DR  Pfam; PF00895; ATP-synt_8; 1.
KW  Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
SQ  SEQUENCE 55 AA; 6443 MW; D02930C2E346925F CRC64;

```

```

RL  Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC  -!- FUNCTION: This is one of the chains of the nonenzymatic component
CC  -!- (CF(0) subunit) of the mitochondrial ATPase complex.
CC  -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC  H(+) (Out).
CC  -!- SUBCELLULAR LOCATION: Membrane-bound.
CC  -!- SIMILARITY: Belongs to the ATPase protein 8 family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  EMBL; AF154851; AAD41389.1; -.
DR  InterPro; IPR001421; ATPase8_mit.
DR  Pfam; PF00895; ATP-synt_8; 1.
KW  Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
SQ  SEQUENCE 55 AA; 6455 MW; 71E430C2E346924A CRC64;

Query Match 47.5%; Score 47; DB 1; Length 55;
Best Local Similarity 83.3%; Pred. No. 4.9;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  3 PWKWPW 8
DB  49 PWNWPW 54

RESULT 14
ATP8_SALFO
ID  ATP8_SALFO      STANDARD;      PRT;      55 AA.
AC  Q9XN35;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  10-OCT-2003 (Rel. 42, Last annotation update)
DE  ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN  MTATP8 OR ATP8 OR ATPASE8
OS  Salvelinus fontinalis (Brook trout) (Brook char).
OC  Mitochondrion.
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC  Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX  NCBI_TaxID=8038;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Doiron S., Blier P.U., Bernatchez L.;
RT  "A comparative analysis of complete sequence of mitochondrial genome
RT  between brook char (Salvelinus fontinalis) and arctic char (S.
RT  alpinus).";
RL  Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC  -!- FUNCTION: This is one of the chains of the nonenzymatic component
CC  -!- (CF(0) subunit) of the mitochondrial ATPase complex.
CC  -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC  H(+) (Out).
CC  -!- SUBCELLULAR LOCATION: Membrane-bound.
CC  -!- SIMILARITY: Belongs to the ATPase protein 8 family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  EMBL; AF154850; AAD41376.1; -.
DR  InterPro; IPR001421; ATPase8_mit.
DR  Pfam; PF00895; ATP-synt_8; 1.
KW  Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
SQ  SEQUENCE 55 AA; 6443 MW; D02930C2E346925F CRC64;

```

Query Match 47.5%; Score 47; DB 1; Length 55;
Best Local Similarity 83.3%; Pred. No. 4.9;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 PWKNEP 8
|||
DB 49 PNNWFW 54

RESULT 15

ATP8_SCYCA
ID ATP8_SCYCA STANDARD; PRT; 55 AA.
AC O79405;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN M1ATP8 OR ATP8 OR ATPASE8.
OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carchariniiformes;
OC Scyliorhinidae; Scyliorhinus.
OX NCBI_taxID=7830;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
EX MEDLINE=98393590; PubMed=9725850;
RA Delarbre C., Spruyt N., Delmarre C., Gallut C., Barriol V.,
RA Janvier P., Laudet V., Gachelin G.;
RT "The complete nucleotide sequence of the mitochondrial DNA of the
RT dogfish, Scyliorhinus canicula.";
RL Genetics 150:331-344(1998).
CC -!- FUNCTION: This is one of the chains of the nonenzymatic component
CC (CF0) subunit of the mitochondrial ATPase complex.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
CC H(+) (out).
CC -!- SUBCELLULAR LOCATION: Membrane-bound.
CC -!- SIMILARITY: Belongs to the ATPase protein 8 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; Y16067; CAA76023.1; -
DR F01304; T11304.
DR InterPro; IPR001421; ATPase8_mit.
DR Pfam; PF00895; ATP-synt_8; 1.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT TRANSMEM 4 24 POTENTIAL.
SQ SEQUENCE 55 AA; 6607 MW; 075956C2A3DF05B9 CRC64;

Query Match 47.5%; Score 47; DB 1; Length 55;
Best Local Similarity 83.3%; Pred. No. 4.9;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 PWKNEP 8
|||
DB 49 PNNWFW 54

Search completed: May 4, 2004, 15:20:17
Job time : 9.21053 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:17:07 ; Search time 14.3684 Seconds
(without alignments)
46.709 Million cell updates/sec

Title: US-09-444-281-35
Perfect score: 91
Sequence: 1 ILKKWPPWPRRK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	100.0	13	3	US-08-915-314-30
2	91	100.0	13	3	US-08-915-314-62
3	91	100.0	13	3	US-08-915-314-63
4	91	100.0	13	3	US-08-915-314-64
5	91	100.0	13	3	US-08-042-071-36
6	91	100.0	13	4	US-08-030-619-95
7	91	100.0	13	4	US-08-030-619-99
8	91	100.0	13	4	US-08-667-486-30
9	91	100.0	13	4	US-08-667-486-62
10	91	100.0	13	4	US-08-667-486-63
11	91	100.0	13	4	US-08-667-486-64
12	91	100.0	14	3	US-08-915-314-57
13	91	100.0	14	4	US-08-030-619-72
14	91	100.0	14	4	US-08-030-619-108
15	91	100.0	14	4	US-08-667-486-57
16	91	100.0	21	3	US-08-915-314-54
17	91	100.0	21	4	US-08-030-619-69
18	91	100.0	21	4	US-08-667-486-54
19	87	95.6	12	3	US-08-915-314-52
20	87	95.6	12	4	US-08-030-619-67
21	87	95.6	12	4	US-08-667-486-52
22	86	94.5	12	3	US-08-915-314-74
23	86	94.5	12	3	US-08-702-054B-5
24	86	94.5	12	4	US-08-030-619-112
25	86	94.5	12	4	US-08-667-486-74
26	86	94.5	13	3	US-08-915-314-58
27	86	94.5	13	4	US-08-030-619-53

28 86 94.5 13 4 US-09-030-619-109 Sequence 109, Appl
29 86 94.5 13 4 US-09-667-486-58 Sequence 58, Appl
30 86 94.5 14 3 US-08-915-314-59 Sequence 59, Appl
31 86 94.5 14 4 US-09-030-619-54 Sequence 54, Appl
32 86 94.5 14 4 US-09-030-619-110 Sequence 110, Appl
33 86 94.5 14 4 US-09-667-486-59 Sequence 59, Appl
34 86 94.5 15 3 US-08-702-054B-40 Sequence 40, Appl
35 86 94.5 20 3 US-08-915-314-55 Sequence 55, Appl
36 86 94.5 20 4 US-09-030-619-51 Sequence 51, Appl
37 86 94.5 20 4 US-09-667-486-55 Sequence 55, Appl
38 86 94.5 21 3 US-08-915-314-56 Sequence 56, Appl
39 86 94.5 21 4 US-09-030-619-52 Sequence 52, Appl
40 86 94.5 21 4 US-09-667-486-56 Sequence 56, Appl
41 85 93.4 12 3 US-08-915-314-69 Sequence 69, Appl
42 85 93.4 12 4 US-09-030-619-73 Sequence 73, Appl
43 85 93.4 12 4 US-09-667-486-69 Sequence 69, Appl
44 85 93.4 13 3 US-08-915-314-38 Sequence 38, Appl
45 85 93.4 13 3 US-08-915-314-45 Sequence 45, Appl

ALIGNMENTS

RESULT 1
US-08-915-314-30
; Sequence 30, Application US/08915314
; Patent No. 6180604
; GENERAL INFORMATION:
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Exlie, Douglas
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,314
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 6180604tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 660081.405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-915-314-30

Query Match 100.0%; Score 91; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPPWPRRK 13
|||||

Db 1 ILKKWPNWPRK 13

RESULT 2

US-08-915-314-62

; Sequence 62, Application US/08915314

; Patent No. 6180604

; GENERAL INFORMATION:

; APPLICANT: Fraser, Janet R.

; APPLICANT: West, Michael H.P.

; APPLICANT: Krieger, Timothy J.

; APPLICANT: Taylor, Robert

; APPLICANT: Erfile, Douglas

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

; TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN

; NUMBER OF SEQUENCES: 90

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/915,314

; FILING DATE: 20-AUG-1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 6180604tenburg Ph.D., Carol

; REGISTRATION NUMBER: 39,317

; REFERENCE/DOCKET NUMBER: 660081.405

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 62:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 13 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; FEATURE:

; NAME/KEY: Modified-site

; LOCATION: 1

; OTHER INFORMATION: /note= "D-Form of Isoleucine"

US-08-915-314-62

Query Match 100.0%; Score 91; DB 3; Length 13;

Best Local Similarity 100.0%; Pred. No. 1.1e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILKKWPNWPRK 13

Db 1 ILKKWPNWPRK 13

RESULT 3

US-08-915-314-63

; Sequence 63, Application US/08915314

; Patent No. 6180604

; GENERAL INFORMATION:

; APPLICANT: Fraser, Janet R.

; APPLICANT: West, Michael H.P.

; APPLICANT: Krieger, Timothy J.

; APPLICANT: Taylor, Robert

; APPLICANT: Erfile, Douglas

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

; TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN

; NUMBER OF SEQUENCES: 90

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/915,314

; FILING DATE: 20-AUG-1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 6180604tenburg Ph.D., Carol

; REGISTRATION NUMBER: 39,317

; REFERENCE/DOCKET NUMBER: 660081.405

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 63:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 13 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; FEATURE:

; NAME/KEY: Modified-site

; LOCATION: 1

; OTHER INFORMATION: /note= "D-Form of Isoleucine"

US-08-915-314-62

Query Match 100.0%; Score 91; DB 3; Length 13;

Best Local Similarity 100.0%; Pred. No. 1.1e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILKKWPNWPRK 13

Db 1 ILKKWPNWPRK 13

RESULT 4

US-08-915-314-64

; Sequence 64, Application US/08915314

; Patent No. 6180604

; GENERAL INFORMATION:

; APPLICANT: Fraser, Janet R.

; APPLICANT: West, Michael H.P.

; APPLICANT: Krieger, Timothy J.

; APPLICANT: Taylor, Robert

; APPLICANT: Erfile, Douglas

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

; TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN

; NUMBER OF SEQUENCES: 90

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/915,314

; FILING DATE: 20-AUG-1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 6180604tenburg Ph.D., Carol

; REGISTRATION NUMBER: 39,317

; REFERENCE/DOCKET NUMBER: 660081.405

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 63:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 13 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; FEATURE:

; NAME/KEY: Modified-site

; LOCATION: 13

; OTHER INFORMATION: /note= "D-Form of lysine"

US-08-915-314-63

Query Match 100.0%; Score 91; DB 3; Length 13;

Best Local Similarity 100.0%; Pred. No. 1.1e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILKKWPNWPRK 13

Db 1 ILKKWPNWPRK 13

```

; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6180604tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 660081.405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "D-Form of Isoleucine"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 13
; OTHER INFORMATION: /note= "D-Form of Lysine"
;
US-08-915-314-64
Query Match 100.0%; Score 91; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWFWPWRRK 13
Db 1 ILKKWFWPWRRK 13

RESULT 5
US-09-042-071-36
; Sequence 36, Application US/09042071
; Patent No. 6294372
; GENERAL INFORMATION:
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; TITLE OF INVENTION: REPLICATION GENES AND GENE PRODUCTS FROM
; TITLE OF INVENTION: SMALL CRYPTIC PLASMIDS AND METHODS FOR CONSTRUCTING
; TITLE OF INVENTION: CONTROLLED-REPLICATION PLASMID VECTORS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/042,071
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McMaisters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 660081.407
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; OTHER INFORMATION: Cationic Peptide Analogue
;
US-09-042-071-36
Query Match 100.0%; Score 91; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWFWPWRRK 13
Db 1 ILKKWFWPWRRK 13

RESULT 6
US-09-030-619-95
; Sequence 95, Application US/09030619B
; Patent No. 6503881
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
;
US-09-030-619-95
Query Match 100.0%; Score 91; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWFWPWRRK 13
Db 1 ILKKWFWPWRRK 13

RESULT 7
US-09-030-619-99
; Sequence 99, Application US/09030619B
; Patent No. 6503881
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 99
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
;
US-09-030-619-99
```

Query Match 100.0%; Score 91; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILKKPWPWRRK 13
| | | | |
Db 1 ILKKPWPWRRK 13

RESULT 8

US-09-667-486-30
; Sequence 30, Application US/09667486
; Patent No. 6538106

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.
West, Michael H.P.
Krieger, Timothy J.
Taylor, Robert
Erfile, Douglas
Erfile, Douglas

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN

NUMBER OF SEQUENCES: 90

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/667,486

FILING DATE: 22-Sep-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/915,314

FILING DATE: 20-AUG-1997

ATTORNEY/AGENT INFORMATION:

NAME: No. 6538106tenburg Ph.D., Carol

REGISTRATION/DOCKET NUMBER: 660081.405

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 30:

US-09-667-486-30

Query Match 100.0%; Score 91; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILKKPWPWRRK 13
| | | | |
Db 1 ILKKPWPWRRK 13

RESULT 9

US-09-667-486-62

; Sequence 62, Application US/09667486
; Patent No. 6538106

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.

West, Michael H.P.

ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/667,486
FILING DATE: 22-Sep-2000
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: NO. 6538106tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
FEATURE:
NAME/KEY: Modified-site
LOCATION: 13
OTHER INFORMATION: /note= "D-Form of Lysine"
SEQUENCE DESCRIPTION: SEQ ID NO: 63:
US-09-667-486-63

Query Match 100.0%; Score 91; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWPWRRK 13
|||||
DB 1 ILKKWPWPWRRK 13

RESULT 11
US-09-667-486-64
Sequence 64, Application US/09667486
Patent No. 6538106
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
West, Michael H.P.
Krieger, Timothy J.
Taylor, Robert
Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/667,486
FILING DATE: 22-Sep-2000
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314

FILING DATE: 20-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: NO. 6538106tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "D-Form of Isoleucine"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 13
OTHER INFORMATION: /note= "D-Form of Lysine"
SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-09-667-486-64

Query Match 100.0%; Score 91; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWPWRRK 13
|||||
DB 1 ILKKWPWPWRRK 13

RESULT 12
US-08-915-314-57
Sequence 57, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: NO. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids

```
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-915-314-57

Query Match      100.0%; Score 91; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWPWRRK 13
   |||||
Db  1 ILKKWPWPWRRK 13

RESULT 13
US-09-030-619-72
; Sequence 72, Application US/09030619B
; Patent No. 6503881
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030.619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-72

Query Match      100.0%; Score 91; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWPWRRK 13
   |||||
Db  1 ILKKWPWPWRRK 13

RESULT 14
US-09-030-619-108
; Sequence 108, Application US/09030619B
; Patent No. 6503881
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030.619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 108
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-030-619-108

; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-108

Query Match      100.0%; Score 91; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWPWRRK 13
   |||||
Db  1 ILKKWPWPWRRK 13

RESULT 15
US-09-667-486-57
; Sequence 57, Application US/09667486
; Patent No. 6538106
; GENERAL INFORMATION:
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/667.486
; FILING DATE: 22-Sep-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/915.314
; FILING DATE: 20-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6538106tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 660081.405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 57:
US-09-667-486-57

Query Match      100.0%; Score 91; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWPWRRK 13
   |||||
Db  1 ILKKWPWPWRRK 13

Search completed: May 4, 2004, 15:23:52
Job time : 14.3684 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:22:18 ; Search time 37.2895 Seconds
(without alignments)
96.635 Million cell updates/sec

Title: US-09-444-281-35
Perfect score: 91
Sequence: 1 ILKKWPPWRRK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 0%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	DB ID	Description
1	91	100.0	13	US-09-030-619-95
2	91	100.0	13	US-09-030-619-99
3	91	100.0	13	US-10-277-232-95
4	91	100.0	13	US-10-277-232-99
5	91	100.0	13	US-10-252-773-3
6	91	100.0	13	US-10-229-368-5
7	91	100.0	13	US-10-229-368-6
8	91	100.0	13	US-10-229-368-56
9	91	100.0	13	US-10-229-368-57
10	91	100.0	13	US-10-229-368-58
11	91	100.0	13	US-10-225-087-5
12	91	100.0	13	US-10-225-087-6
13	91	100.0	13	US-10-225-087-52
14	91	100.0	13	US-10-225-087-53
15	91	100.0	13	US-10-225-087-54

16	91	100.0	13	US-10-277-233-95
17	91	100.0	13	US-10-277-233-99
18	91	100.0	13	US-10-351-985-30
19	91	100.0	13	US-10-351-985-62
20	91	100.0	13	US-10-351-985-63
21	91	100.0	13	US-10-351-985-64
22	91	100.0	13	US-10-403-104-1
23	91	100.0	14	US-09-030-619-72
24	91	100.0	14	US-09-030-619-108
25	91	100.0	14	US-10-277-232-72
26	91	100.0	14	US-10-277-232-108
27	91	100.0	14	US-10-229-368-48
28	91	100.0	14	US-10-225-087-45
29	91	100.0	14	US-10-277-233-72
30	91	100.0	14	US-10-277-233-108
31	91	100.0	14	US-10-351-985-57
32	91	100.0	21	US-09-030-619-69
33	91	100.0	21	US-10-277-232-69
34	91	100.0	21	US-10-229-368-43
35	91	100.0	21	US-10-225-087-40
36	91	100.0	21	US-10-277-233-69
37	91	100.0	21	US-10-351-985-54
38	87	95.6	12	US-09-030-619-67
39	87	95.6	12	US-10-277-232-67
40	87	95.6	12	US-10-229-368-41
41	87	95.6	12	US-10-225-087-38
42	87	95.6	12	US-10-277-233-67
43	87	95.6	12	US-10-351-985-52
44	86	94.5	12	US-09-030-619-112
45	86	94.5	12	US-10-277-232-112

ALIGNMENTS

RESULT 1

US-09-030-619-95
; Sequence 95, Application US/09030619B
; Patent No. US2002035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Efile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-95

Query Match 100.0%; Score 91; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILKKWPPWRRK 13
| | | | | | | | | | | | |
Db 1 ILKKWPPWRRK 13

RESULT 2
US-09-030-619-99

; Sequence 99, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieser, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Exile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 99
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
; US-09-030-619-99

Query Match 100.0%; Score 91; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWPWRRK 13
| | | | | | | | | | | | | |
DB 1 ILKKWPWPWRRK 13

RESULT 3
US-10-277-232-95
; Sequence 95, Application US/10277232
; Publication No. US20030211537A1
; GENERAL INFORMATION:
; APPLICANT: Krieser, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Exile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406C1
; CURRENT APPLICATION NUMBER: US/10/277,232
; CURRENT FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
; US-10-277-232-95

Query Match 100.0%; Score 91; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWPWRRK 13
| | | | | | | | | | | | | |
DB 1 ILKKWPWPWRRK 13

RESULT 4
US-10-277-232-99
; Sequence 99, Application US/10277232

; Publication No. US20030211537A1
; GENERAL INFORMATION:
; APPLICANT: Krieser, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Exile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406C1
; CURRENT APPLICATION NUMBER: US/10/277,232
; CURRENT FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 99
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
; US-10-277-232-99

Query Match 100.0%; Score 91; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWPWRRK 13
| | | | | | | | | | | | | |
DB 1 ILKKWPWPWRRK 13

RESULT 5
US-10-252-773-3
; Sequence 3, Application US/10252773
; Publication No. US2003013183A1
; GENERAL INFORMATION:
; APPLICANT: EVERETT, NICHOLAS P.
; APPLICANT: LI, QUNIGSHUN
; APPLICANT: LAWRENCE, CHRISTOPHER
; APPLICANT: DAVIES, MAELOR H.
; TITLE OF INVENTION: PEPTIDES WITH ENHANCED STABILITY TO PROTEASE
; TITLE OF INVENTION: DEGRADATION
; FILE REFERENCE: INTERLINK 3.0-003
; CURRENT APPLICATION NUMBER: US/10/252,773
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 60/106,373
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106,573
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: antimicrobial peptide
; US-10-252-773-3

Query Match 100.0%; Score 91; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWPWRRK 13
| | | | | | | | | | | | | |
DB 1 ILKKWPWPWRRK 13

RESULT 6
US-10-229-368-5

; Sequence 5, Application US/10229368
; Publication No. US20030148945A1
; GENERAL INFORMATION:
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Pawlak, Sonia K.
; APPLICANT: Rubinchik, Evelina
; APPLICANT: Cameron, Dale
; APPLICANT: Guarna, Maria Marta
; TITLE OF INVENTION: ANTIMICROBIAL AND ANTI-INFLAMMATORY
; FILE REFERENCE: 660081.418
; CURRENT APPLICATION NUMBER: US/10/229,368
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin peptide analogs
US-10-229-368-5

Query Match 100.0%; Score 91; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILKKWPWPWRRK 13
||| ||||| |||||
Db 1 ILKKWPWPWRRK 13

RESULT 7

US-10-229-368-6
; Sequence 6, Application US/10229368
; Publication No. US20030148945A1
; GENERAL INFORMATION:
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Pawlak, Sonia K.
; APPLICANT: Rubinchik, Evelina
; APPLICANT: Cameron, Dale
; APPLICANT: Guarna, Maria Marta
; TITLE OF INVENTION: ANTIMICROBIAL AND ANTI-INFLAMMATORY
; FILE REFERENCE: 660081.418
; CURRENT APPLICATION NUMBER: US/10/229,368
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin peptide analogs
US-10-229-368-6

Query Match 100.0%; Score 91; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILKKWPWPWRRK 13
||| ||||| |||||
Db 1 ILKKWPWPWRRK 13

RESULT 8

US-10-229-368-56
; Sequence 56, Application US/10229368
; Publication No. US20030148945A1
; GENERAL INFORMATION:
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Pawlak, Sonia K.

; APPLICANT: Rubinchik, Evelina
; APPLICANT: Cameron, Dale
; APPLICANT: Guarna, Maria Marta
; TITLE OF INVENTION: ANTIMICROBIAL AND ANTI-INFLAMMATORY
; FILE REFERENCE: 660081.418
; CURRENT APPLICATION NUMBER: US/10/229,368
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin peptide analogs
; NAME/KEY: MOD_RES
; LOCATION: (1)...(1)
; OTHER INFORMATION: D-Isoleucine
US-10-229-368-56

Query Match 100.0%; Score 91; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILKKWPWPWRRK 13
||| ||||| |||||
Db 1 ILKKWPWPWRRK 13

RESULT 9

US-10-229-368-57
; Sequence 57, Application US/10229368
; Publication No. US20030148945A1
; GENERAL INFORMATION:
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Pawlak, Sonia K.
; APPLICANT: Rubinchik, Evelina
; APPLICANT: Cameron, Dale
; APPLICANT: Guarna, Maria Marta
; TITLE OF INVENTION: ANTIMICROBIAL AND ANTI-INFLAMMATORY
; FILE REFERENCE: 660081.418
; CURRENT APPLICATION NUMBER: US/10/229,368
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin peptide analogs
; NAME/KEY: MOD_RES
; LOCATION: (13)...(13)
; OTHER INFORMATION: D-Lysine
US-10-229-368-57

Query Match 100.0%; Score 91; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILKKWPWPWRRK 13
||| ||||| |||||
Db 1 ILKKWPWPWRRK 13

RESULT 10

US-10-229-368-58
; Sequence 58, Application US/10229368
; Publication No. US20030148945A1

```

; GENERAL INFORMATION:
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Pawlak, Sonia K.
; APPLICANT: Rubinchik, Evelina
; APPLICANT: Cameron, Dale
; APPLICANT: Guarna, Maria Marta
; TITLE OF INVENTION: ANTIMICROBIAL AND ANTI-INFLAMMATORY
; TITLE OF INVENTION: PEPTIDES
; FILE REFERENCE: 660081.418
; CURRENT APPLICATION NUMBER: US/10/229,368
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin peptide analogs
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (1)...(1)
; OTHER INFORMATION: D-Isoleucine
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (13)...(13)
; OTHER INFORMATION: D-Lysine
US-10-229-368-58

```

```

Query Match      100.0%; Score 91; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 ILKKWPWPWRRK 13
Db 1 ILKKWPWPWRRK 13

```

```

RESULT 11
US-10-225-087-5
; Sequence 5, Application US/10225087
; Publication No. US20030171281A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Fraser, Janet R.
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES AND
; TITLE OF INVENTION: FORMULATIONS THEREOF
; FILE REFERENCE: 660081.417
; CURRENT APPLICATION NUMBER: US/10/225,087
; CURRENT FILING DATE: 2003-01-10
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin analog
US-10-225-087-5

```

```

Query Match      100.0%; Score 91; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 ILKKWPWPWRRK 13
Db 1 ILKKWPWPWRRK 13

```

```

RESULT 12
US-10-225-087-6
; Sequence 6, Application US/10225087

```

```

; Publication No. US20030171281A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Fraser, Janet R.
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES AND
; TITLE OF INVENTION: FORMULATIONS THEREOF
; FILE REFERENCE: 660081.417
; CURRENT APPLICATION NUMBER: US/10/225,087
; CURRENT FILING DATE: 2003-01-10
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin analog
US-10-225-087-6

```

```

Query Match      100.0%; Score 91; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 ILKKWPWPWRRK 13
Db 1 ILKKWPWPWRRK 13

```

```

RESULT 13
US-10-225-087-52
; Sequence 52, Application US/10225087
; Publication No. US20030171281A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Fraser, Janet R.
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES AND
; TITLE OF INVENTION: FORMULATIONS THEREOF
; FILE REFERENCE: 660081.417
; CURRENT APPLICATION NUMBER: US/10/225,087
; CURRENT FILING DATE: 2003-01-10
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin analog
US-10-225-087-52

```

```

Query Match      100.0%; Score 91; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 ILKKWPWPWRRK 13
Db 1 ILKKWPWPWRRK 13

```

```

RESULT 14
US-10-225-087-53
; Sequence 53, Application US/10225087
; Publication No. US20030171281A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Fraser, Janet R.
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES AND
; TITLE OF INVENTION: FORMULATIONS THEREOF
; FILE REFERENCE: 660081.417
; CURRENT APPLICATION NUMBER: US/10/225,087

```

```
; CURRENT FILING DATE: 2003-01-10
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin analog
US-10-225-087-53

Query Match      100.0%; Score 91; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWWPWRK 13
Db 1 ILKKWPWWPWRK 13

RESULT 15
US-10-225-087-54
; Sequence 54, Application US/10225087
; Publication No. US20030171281A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Frazer, Janet R.
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES AND
; FILE REFERENCE: 660081.417
; CURRENT APPLICATION NUMBER: US/10/225,087
; CURRENT FILING DATE: 2003-01-10
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin analog
US-10-225-087-54

Query Match      100.0%; Score 91; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWWPWRK 13
Db 1 ILKKWPWWPWRK 13

Search completed: May 4, 2004, 15:35:35
Job time : 37.2895 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:08:11 ; Search time 49.6053 Seconds
(without alignments)
74.047 Million cell updates/sec

Title: US-09-444-281-35

Perfect score: 91

Sequence: 1 ILKKWPPWRRK 13

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1980s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	100.0	13	2 AAW12873	Aaw12873 Antimicro
2	91	100.0	13	2 AAW71690	Aaw71690 Cationic
3	91	100.0	13	2 AAW66378	Aaw66378 Cationic
4	91	100.0	13	2 AAY24609	Aay24609 Indolicid
5	91	100.0	13	2 AAY94495	Aay94495 MBI-11 pe
6	91	100.0	13	3 AAY91818	Aay91818 Amino aci
7	91	100.0	13	3 AAY91774	Aay91774 Amino aci
8	91	100.0	13	3 AAY91773	Aay91773 Amino aci
9	91	100.0	13	3 AAY91820	Aay91820 Amino aci
10	91	100.0	13	3 AAY91819	Aay91819 Amino aci
11	91	100.0	13	3 AAY92795	Aay92795 Indolicid
12	91	100.0	13	5 ABB81254	Abb81254 Cpl1-NH2
13	91	100.0	13	6 ADA00554	Ada00554 Antimicro
14	91	100.0	13	6 ADA00506	Ada00506 Antimicro
15	91	100.0	13	6 ADA00555	Ada00555 Antimicro
16	91	100.0	13	6 ADA00553	Ada00553 Antimicro
17	91	100.0	13	6 ADA00507	Ada00507 Antimicro
18	91	100.0	13	7 ADC73323	Adc73323 Bovine an
19	91	100.0	13	7 ADC98853	Adc98853 Synthetic
20	91	100.0	13	7 ADC98905	Adc98905 Synthetic
21	91	100.0	13	7 ADC98906	Adc98906 Synthetic
22	91	100.0	13	7 ADC98854	Adc98854 Synthetic
23	91	100.0	13	7 ADC98904	Adc98904 Synthetic
24	91	100.0	14	2 AAY24583	Aay24583 Indolicid
25	91	100.0	14	3 AAY91811	Aay91811 Amino aci

26	91	100.0	14	6 ADA00546	Ada00546 Antimicro
27	91	100.0	14	7 ADC98896	Adc98896 Synthetic
28	91	100.0	21	2 AAY24582	Aay24582 Indolicid
29	91	100.0	21	3 AAY91806	Aay91806 Amino aci
30	91	100.0	21	6 ADA00541	Ada00541 Antimicro
31	91	100.0	21	7 ADC98891	Adc98891 Synthetic
32	87	95.6	12	2 AAY24580	Aay24580 Indolicid
33	87	95.6	12	3 AAY91804	Aay91804 Amino aci
34	87	95.6	12	6 ADA00539	Ada00539 Antimicro
35	87	95.6	12	7 ADC98889	Adc98889 Synthetic
36	86	94.5	12	2 AAW12877	Aaw12877 Antimicro
37	86	94.5	12	2 AAY24615	Aay24615 Indolicid
38	86	94.5	12	3 AAY91833	Aay91833 Amino aci
39	86	94.5	12	6 ADA00578	Ada00578 Antimicro
40	86	94.5	12	7 ADC98935	Adc98935 Synthetic
41	86	94.5	13	2 AAY24572	Aay24572 Indolicid
42	86	94.5	13	3 AAY91812	Aay91812 Amino aci
43	86	94.5	13	6 ADA00547	Ada00547 Antimicro
44	86	94.5	13	7 ADC98897	Adc98897 Synthetic
45	86	94.5	14	2 AAY24573	Aay24573 Indolicid

ALIGNMENTS

RESULT 1

AAW12873
ID AAW12873 standard; peptide; 13 AA.

XX AC AAW12873;
XX AC
XX 10-DEC-1997 (first entry)
XX DT
XX DE Antimicrobial cationic peptide CP-11.
XX KW Bacterial; viral; antitumour; food; preservative; inhibitor; growth;
XX KW Bacterium; yeast; endotoxaemia; sepsis; antibiotic; fungal; antiviral;
XX KW Candida albicans; steriliant; Salmonella; Yersina; Shigella.

OS Synthetic.

XX WO9708199-A2.

XX 06-MAR-1997.

XX 23-AUG-1996; 96WO-IB0000996.

XX 23-AUG-1995; 95US-0002687P.

XX (UYBR-) UNIV BRITISH COLUMBIA.

XX Falla TU, Hancock REW, Gough M;

XX WPI; 1997-179179/16.

XX Cationic peptide(s) having anti-microbial activity - used for the inhibition of bacterial and viral growth, as an antitumour agent, and as a food preservative.

XX Claim 2; Page 65; 89pp; English.

XX The present sequence represents a specifically claimed novel isolated cationic peptide which has antimicrobial activity. The amino acid sequence of antimicrobial cationic peptides (including the present sequence) is selected from: XIX1ProX2X3X2Pro(X2X2Pro)nx2X3(X5)o; CC XIX1ProX2X3X4(X5)rProX2X3X3; XIX1X3(ProTrp)uX3X2X5X2(X5)o; CC XIX1X3X3X2Pro(X2X2Pro)nx2(X5)m; where m = 1-5; n = 1-2; o = 2-5; r = 0-8; CC u = 0-1; X1 = Ile, Leu, Val, Phe, Tyr, Trp or Met; X2 = Trp or Phe; X3 = Arg or Lys; X4 = Trp or Lys; and X5 = Phe, Trp, Arg, Lys or Pro. The peptides are preferably amidated or carboxymethylated. The peptides may be used in methods for inhibiting the growth of a bacterium or yeast, or for inhibiting an endotoxaemia or sepsis associated disorder in a subject. The peptides have a broad activity against antibiotic resistant

CC bacteria, combined with activity against the medically important fungus
 CC Candida albicans. In addition, the peptides are useful as antitumour
 CC agents and/or antiviral agents. The peptides may be used as sterilants or
 CC preservatives of materials susceptible to microbial or viral
 CC contamination, e.g. in processed foods to inhibit Salmonella, Yersinia and
 CC Shigella. The peptides are compact and tend to have a unique polypeptide
 CC type II extended helix structure that permits them to span the membrane
 CC with relatively few amino acids. The peptides possess the ability to work
 CC synergistically with antibiotics, and in addition, some of them possess
 CC anti-endotoxin activity. N.B. The present sequence represents SEQ ID NO:1
 CC in the claims and examples of the specification, but differs slightly
 CC from the SEQ ID NO:1 in the sequence listing on page 51 of the
 CC specification (see AAW21719)

XX
 XX
 SQ Sequence 13 AA;

Query Match 100.0%; Score 91; DB 2; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWFWPWRRK 13
 |||||
 Db 1 ILKKWFWPWRRK 13

RESULT 2
 AAW71690
 ID AAW71690 standard; peptide; 13 AA.

XX
 AC AAW71690;

DT 11-JAN-1999 (first entry)

DE Cationic peptide MB111 (MW 1879).

KW MB111; cationic peptide; plasmid pKLI; small cryptic plasmid;
 KW replication; RepA; vector; RAMP.

OS Synthetic.

PN WO9841636-A2.

PD 24-SEP-1998.

PF 16-MAR-1998; 98WO-CA000214.

PR 14-MAR-1997; 97US-0040722P.

XX (BURI/) BURIAN J.
 XX (KAYW/) KAY W W.

PI Burian J, Kay WW;

XX WPI; 1998-531571/45.

XX Increasing plasmid copy number in a cell with the repA gene product - and
 PT an small cryptic plasmid ori sequence, useful for high level expression
 PT of e.g. cytokines, antigens or therapeutic proteins.

XX Example 13; Page 54; 82pp; English.

XX MB111 is a small (mol.wt. 1879) cationic peptide. DNA encoding MB111 has
 CC been incorporated into vector pR2h-B1, in which the replication leader
 CC (R21) sequence of RepA (see also AAW71686) is joined to 2 Hero peptides
 CC (see also AAW71692), to provide a vector for expression of MB111 in host
 CC cells. The invention provides controlled replication plasmid vectors
 CC (RAMP vectors) comprising a replication origin of a small cryptic plasmid
 CC and a gene encoding RepA. The vectors can reach very high levels of
 CC plasmid replication, but are not lethal to the host cell, and can be used
 CC to direct the high level expression of e.g. cytokines, antigens and
 CC therapeutic proteins

XX Sequence 13 AA;

Query Match 100.0%; Score 91; DB 2; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWFWPWRRK 13
 |||||
 Db 1 ILKKWFWPWRRK 13

RESULT 3
 AAW66378
 ID AAW66378 standard; peptide; 13 AA.

XX
 AC AAW66378;

DT 12-JAN-1999 (first entry)

DE Cationic peptide of claim 15 #5.

KW Indolicidin analogue; resistance; cationic peptide; antibiotic;
 KW bacterial infection; tolerance; antibacterial; microorganism; bacteria;
 KW fungus; parasite; virus.

OS Synthetic.

PN WO9840401-A2.

PD 17-SEP-1998.

PF 10-MAR-1998; 98WO-CA000190.

PR 10-MAR-1997; 97US-0040649P.

PR 20-AUG-1997; 97US-00915314.

PR 26-SEP-1997; 97US-0060099P.

PR 25-FEB-1998; 98US-00030619.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Fraser JR, West MHP, McNicol PJ;

XX WPI; 1998-520800/44.

XX New indolicidin peptide analogues - useful for, e.g. enhancing activity
 PT of antibiotic or overcoming tolerance, acquired resistance or inherent
 PT resistance of microorganisms.

XX Claim 15; Page 93; 105pp; English.

XX The present sequence represents a specifically claimed cationic peptide
 CC from the present invention. The present invention describes compositions
 CC and methods for treating infection, especially bacterial infections. The
 CC compositions and methods use cationic peptides in combination with an
 CC antibiotic agent which are then administered to a patient to enhance the
 CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
 CC acquired resistance; and (c) inherent resistance. The combinations of
 CC antibiotics and cationic peptides can provide synergistic activity
 CC against a microorganism that is tolerant, inherently resistant, or has
 CC acquired resistance to an antibiotic agent. They can be used for killing
 CC e.g. bacteria, fungi, parasites and viruses

XX Sequence 13 AA;

Query Match 100.0%; Score 91; DB 2; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWFWPWRRK 13
 |||||
 Db 1 ILKKWFWPWRRK 13

RESULT 4

RAY24609
ID AAY24609 standard; peptide; 13 AA.

AC AAY24609;

DT 18-AUG-1999 (first entry)

DE Indolicidin analogue #61.

XX Indolicidin; bacterial infection; photo-oxidised solubiliser;
KW antimicrobial; antibiotic; antiarrhythmic; surface disinfectant; additive;
KW shampoo; soap; insecticide; herbicide; preservative; food;
KW technical material.

OS Synthetic.

XX WO9807745-A2.

XX 26-FEB-1998.

XX 21-AUG-1997; 97WO-US014779.

XX 21-AUG-1996; 96US-0024754P.

PR 13-JAN-1997; 97US-0034949P.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Fraser JR, West MH, Krieger TJ, Taylor R, Erfle D;

XX WPI; 1998-169090/15.

XX New indolicidin analogues with antimicrobial activity and related nucleic
PT acid - vectors, transformed cells and antibodies, also conjugates with
PT polyoxyalkylene glycol and fatty acid to reduce toxicity, useful
PT therapeutically, as disinfectants etc.

XX Example 1; Page 32; 129pp; English.

XX AAY24549 to AAY24615 represent indolicidin analogues of formulae (I) -
CC (VII) containing up to 25 amino acids (aa): RXZXXZXB (I), BXZXXZXB (II),
CC BBXZXXZXB (III), BXZXXZXBBA(AA)NMLBBAGS (IV), BXZXXZXB(AA)NM (V),
CC LBXZXXZBXZXRK (VI), LKXZXXZXRK (VII) and BBXZXXZXB (VIII). Where Z =
CC P or V; X = hydrophobic residue, preferably W; B = basic aa, preferably R
CC or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V; in (VIII) at
CC least 2 X = F or Y. The analogues are used to treat infections caused by
CC bacteria (Gram positive or negative, or anaerobic); fungi (yeast or
CC moulds); parasites (protozoa, nematodes, cestodes or trematodes) or
CC viruses. Typical of very many pathogens that can be controlled are
CC Leishmania, trypanosoma, Ascaris lumbricoides, Fasciola hepatica,
CC Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus aureus,
CC Listeria, Clostridium, rotavirus and papilloma virus. Compounds derived
CC from the analogues may be used similarly; the compounds may also be
CC prepared from antibiotics or antiarrhythmic agents. The analogues may be
CC used therapeutically or to coat medical devices; also they are useful as
CC surface disinfectants, as additives to shampoo or soaps, as insecticides
CC or herbicides, or as preservatives for foods and technical materials. The
CC analogues are administered by injection, lavage, orally or topically,
CC generally at 0.1-50 mg/kg. These analogues have a broader spectrum of
CC activity than indolicidin and modification as compounds reduces their
CC toxicity

XX Sequence 13 AA;

Query Match 100.0%; Score 91; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKQWFWPWRK 13

Db 1 ILKQWFWPWRK 13

RESULT 5

AAY94495
ID AAY94495 standard; peptide; 13 AA.

XX AAY94495;

XX 20-SEP-2000 (first entry)

XX MBI-11 peptide derived from indolicidin.

XX Cellulose binding domain; CBD; cationic peptide; MBI-11; indolicidin;
KW bovine.

OS Bos taurus.

XX WO2000031279-A2.

XX 02-JUN-2000.

XX 19-NOV-1999; 99WO-CA001107.

XX 20-NOV-1998; 98US-0109218P.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Burian J, Bartfeld D;

XX WPI; 2000-400086/34.

XX Multi-domain fusion protein expression cassette used for high yield
PT stable production of foreign peptide gene products.

XX Disclosure; Page 24; 73pp; English.

XX A novel method allows the efficient production of cationic peptides in:
CC recombinant host cells. The method involves construction of a multi-
CC domain fusion protein expression cassette comprising a promoter and a
CC nucleic acid molecule expressed as an insoluble protein. The inclusion of
CC anionic peptide sequences in the linker sequences neutralises the fusion
CC positive charge of the cationic peptide so that the charge of the fusion
CC protein is controlled. This cassette allows high yield, stable production
CC of the cationic peptide. Cationic peptides such as bovine indolicidin may
CC be used as antimicrobial agents. The present sequence is the MBI-11
CC peptide. MBI-11 is a cationic peptide derived from modifications of
CC indolicidin

XX Sequence 13 AA;

Query Match 100.0%; Score 91; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKQWFWPWRK 13

Db 1 ILKQWFWPWRK 13

RESULT 6

RAY91818

ID AAY91818 standard; peptide; 13 AA.

XX AAY91818;

XX 06-JUN-2000 (first entry)

XX Amino acid sequence of cationic peptide MBI 11E1CN.

XX Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
KW multidrug resistance.

OS Synthetic.

XX

Best Local Similarity 100.0%; Pred. No. 2.1e-06; Mismatches 0; Indels 0; Gaps 0;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWPWRRK 13
 DB 1 ILKKWPWPWRRK 13

RESULT 9
 AAY91820
 ID AAY91820 standard; peptide; 13 AA.
 XX AC
 XX DT
 XX 06-JUN-2000 (first entry)
 DE Amino acid sequence of cationic peptide MBI 11B3CN.
 XX KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.
 XX OS Synthetic.
 XX PN WO9965506-A2.
 XX PD 23-DEC-1999.
 XX PF 14-JUN-1999; 99WO-CA000552.
 XX PR 12-JUN-1998; 98US-00096541.
 XX PA (MICR-) MICROLOGIX BIOTECH INC.
 XX PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
 XX WPI; 2000-223549/19.
 XX DR Novel pharmaceutical composition containing optionally activated
 PT polyoxyalkylene-modified cationic peptides, useful for treating tumors.
 XX PS Disclosure; Page 15; 94pp; English.
 XX CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxyalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon
 XX SQ Sequence 13 AA;

Query Match 100.0%; Score 91; DB 3; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWPWRRK 13
 DB 1 ILKKWPWPWRRK 13

RESULT 10
 AAY91819
 ID AAY91819 standard; peptide; 13 AA.
 XX AC
 XX DT
 XX 06-JUN-2000 (first entry)

DE Amino acid sequence of cationic peptide MBI 11B2CN.
 XX KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.
 XX OS Synthetic.
 XX PN WO9965506-A2.
 XX PD 23-DEC-1999.
 XX PF 14-JUN-1999; 99WO-CA000552.
 XX PR 12-JUN-1998; 98US-00096541.
 XX PA (MICR-) MICROLOGIX BIOTECH INC.
 XX PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
 XX WPI; 2000-223549/19.
 XX DR Novel pharmaceutical composition containing optionally activated
 PT polyoxyalkylene-modified cationic peptides, useful for treating tumors.
 XX PS Disclosure; Page 15; 94pp; English.
 XX CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxyalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon
 XX SQ Sequence 13 AA;

Query Match 100.0%; Score 91; DB 3; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPWPWRRK 13
 DB 1 ILKKWPWPWRRK 13

RESULT 11
 AAY92795
 ID AAY92795 standard; peptide; 13 AA.
 XX AC
 XX DT
 XX 29-AUG-2000 (first entry)
 DE Indolicidin analogue, CP-11.
 XX KW Magainin; antimicrobial; transgenic plant; protease degradation; Rev4;
 KW indolicidin; protein production; reverse peptide.
 XX OS Synthetic.
 XX PN WO200026344-A1.
 XX PD 11-MAY-2000.
 XX PF 29-OCT-1999; 99WO-US025561.
 XX PR 30-OCT-1998; 98US-0106373P.
 XX PR 02-NOV-1998; 98US-0106537P.

PA (INTE-) INTERLINK BIOTECHNOLOGIES LLC.
XX (KENT) UNIV KENTUCKY RES FOUND.
PI Everett NP, Li Q, Lawrence C, Davies MH;
XX WPI; 2000-365597/31.
DR
XX
XX
PT Polypeptides for reducing proteolytic degradation of proteins
PT administered to, or produced by a plant comprise indolicin or its
PT functional equivalents.
XX
XX
PS Disclosure; Page 4; 50pp; English.
XX
XX Indolicin is a potent antimicrobial tridecapeptide, originally purified
CC from cytoplasmic granules of bovine neutrophils. CP-11 is an analogue,
CC which has better activity against E. coli, Pseudomonas aeruginosa and
CC Candida albicans, but reduced activity against Staphylococcus aureus. A
CC reverse peptide, Rev4 (AAV92796) of indolicin was found to have
CC increased stability against plant protease degradation. Expression of
CC antimicrobial peptides in transgenic plants suffers a major limitation in
CC that the foreign peptides are susceptible to rapid degradation by
CC proteases. The invention concerns reducing the extent of protease
CC degradation of a protein applied to, or produced by a plant by
CC administering indolicin, Rev4 or a functional equivalent to the plant.
CC Transgenic plants expressing indolicin and Rev4 are useful for
CC production of the antimicrobial peptides. Compositions containing
CC indolicin and Rev4 are also useful for production of agronomically
CC important proteins in plants
XX
SQ Sequence 13 AA;
Query Match 100.0%; Score 91; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILKKPWPWPWRK 13
DB 1 ILKKPWPWPWRK 13
RESULT 12
ID ABB81254
XX ABB81254 standard; peptide; 13 AA.
XX
AC ABB81254;
DT 20-AUG-2002 (first entry)
XX
DE CP11-NH2 antibacterial peptide SEQ ID NO:21.
XX
XX Antibacterial; glycopeptide; peptidic membrane associating element;
KW bacterial infection; vancomycin; peptidoglycan biosynthesis inhibition;
KW antibiotic.
XX
XX Synthetic.
OS
XX Key Location/Qualifiers
FH Modified-site 13
FT /note= "amidated"
FT
XX WC200236612-A1.
XX
XX 10-MAY-2002.
PD
XX 02-NOV-2001; 2001WO-GB004867.
XX
XX 03-NOV-2000; 2000GB-00026924.
XX
XX (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.
PA (ADPR-) ADPROTECH LTD.
XX
XX Cooper MA, Betley JR;
XX

DR WPI; 2002-471498/50.
XX
XX Antibacterial compound, useful for the treatment of a bacterial infection
PT by e.g. gram positive or negative bacteria, comprises a conjugate of
PT glycopeptide and peptidic membrane-associating element.
XX
XX Disclosure; Page 21; 64pp; English.
XX
XX The present invention describes an antibacterial compound (I), comprising
CC a conjugates of glycopeptide and peptidic membrane-associating elements.
CC (I) comprises the formula V-L-W-X, where: V = a glycopeptide moiety that
CC inhibits peptidoglycan biosynthesis in bacteria; L = a linking group; W =
CC a peptidic membrane-associating element; and X = H or a membrane-
CC insertive element. Also described: (1) a method of treating or preventing
CC a bacterial infection, comprising the administration of (I); and (2) use
CC of (I) in the manufacture of a medicament for the treatment or prevention
CC for a bacterial infection. (I) are used in the manufacture of a medicament
CC for the treatment or prophylaxis of a bacterial infection in a human or
CC animal body, including both the gram positive and gram negative bacteria
CC including Mycobacterium sp., Vibrio sp., Neisseria sp., Borrelia sp., Klebsiella
CC sp., Hemophilus sp., Clostridium sp., Pseudomonas sp., Actinomyces sp.,
CC Staphylococcus sp., Enterococcus sp., Escherichia sp.,
CC Pneumococcus sp. or Salmonella sp., particularly antibiotic resistant
CC bacterial strains. (I) are also useful as wound treatment agents to
CC prevent adhesion of bacteria to matrix proteins, especially fibronectin,
CC exposed in wound tissue; and for prophylactic use in dental treatment as
CC an alternative to, or in conjunction with, antibiotic prophylaxis. (I)
CC has stronger binding to bacterial membranes which have a higher
CC proportion of acidic phospholipids than the eukaryotic organisms, also
CC having a higher proportion of membrane associated biosynthetic proteins.
CC Vancomycin shows an enhanced antimicrobial activity upon derivatisation
CC with (I) and is effective to treat the antibiotic resistant bacterial
CC strains. ABB81234 to ABB81272 represent peptides given in the
CC exemplification of the present invention
XX
SQ Sequence 13 AA;
Query Match 100.0%; Score 91; DB 5; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILKKPWPWPWRK 13
DB 1 ILKKPWPWPWRK 13
RESULT 13
ID ADA00554
XX ADA00554 standard; peptide; 13 AA.
XX
AC ADA00554;
DT 06-NOV-2003 (first entry)
XX
XX Antimicrobial cationic peptide 11E2CN.
DE
XX antimicrobial; cationic; viscosity-increasing agent; solvent; buffer;
KW antibacterial; virucide; antiinflammatory; fungicide; protozoacide;
KW parasiticide; vulnerary; dermatological; herbicide; insecticide;
KW infection; systemic infection; sepsis; acne; disinfectant; herbicide;
KW insecticide; silicone sealant.
XX
XX Synthetic.
OS
XX Key Location/Qualifiers
FH Modified-site 13
FT /label= amidated
FT
XX WO2003015809-A2.
XX
XX 27-FEB-2003.
XX
XX 21-AUG-2002; 2002WO-US026525.
XX

XX 21-AUG-2001; 2001US-0314232P.
 PR 20-AUG-2002; 2002US-00225087.
 XX (MICR-) MICROLOGIX BIOTECH INC.
 XX Krieger TJ, McNicol PJ, Fraser JR;
 XX WPI; 2003-332767/31.
 XX Composition containing stabilized antimicrobial cationic protein, useful
 PT for treating infections, particularly where associated with in-dwelling
 PT devices.
 XX Example 1; Page 48; 90pp; English.
 PS The present invention describes a composition (A) comprising an
 XX antimicrobial cationic peptide (I), a viscosity-increasing agent (II) and
 CC a solvent (III). Also described is a composition comprising (I), buffer
 CC (IV) and (III). (I) has antibacterial, virucide, antiinflammatory,
 CC fungicide, protozoacide, parasiticide, vulnery, dermatological,
 CC herbicide and insecticide activities. (A) can be used to reduce the
 CC population of microflora (eukaryotes, prokaryotes or viruses) at a target
 CC site, particularly for treatment or prevention of infections. They can be
 CC used to treat a wide range of systemic infections (e.g. sepsis) and for
 CC topical treatment of wounds, but most especially can be used: (i) at
 CC sites where medical devices have been, or will be, inserted into the body
 CC (alternatively, they are used to treat the devices); and (ii) at sites on
 CC the skin (particularly for treating acne) or the mucosa. The devices
 CC treated are especially central venous, vascular dialysis, pulmonary
 CC artery, peritoneal dialysis or umbilical catheters. They may also be used
 CC as surface disinfectants; for treatment of clothing and air filters; in
 CC cosmetics and soaps; as herbicides and insecticides; in building
 CC materials (e.g. silicone sealants) and in processing animal products,
 CC e.g. hides. The present sequence represents an antimicrobial cationic
 CC peptide, which is used in the exemplification of the present invention.
 XX Sequence 13 AA;
 SQ

Query Match 100.0%; Score 91; DB 6; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWFWWPRRK 13
 |||||
 DB 1 ILKKWFWWPRRK 13
 |||||

RESULT 14
 ADA00506
 ID ADA00506 standard; peptide; 13 AA.
 XX
 AC ADA00506;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Antimicrobial cationic peptide 11.
 XX
 KW antimicrobial; cationic; viscosity-increasing agent; solvent; buffer;
 KW antibacterial; virucide; antiinflammatory; fungicide; protozoacide;
 KW parasiticide; vulnery; dermatological; herbicide; insecticide;
 KW infection; systemic infection; sepsis; acne; disinfectant; herbicide;
 KW insecticide; silicone sealant.
 XX
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT Modified-site 13
 FT /label= amidated
 XX
 PN WO2003015809-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 21-AUG-2002; 2002WO-US026525.
 XX
 PR 21-AUG-2001; 2001US-0314232P.

PR 20-AUG-2002; 2002US-00225087.
 XX (MICR-) MICROLOGIX BIOTECH INC.
 XX Krieger TJ, McNicol PJ, Fraser JR;
 XX WPI; 2003-332767/31.
 XX Composition containing stabilized antimicrobial cationic protein, useful
 PT for treating infections, particularly where associated with in-dwelling
 PT devices.
 XX Example 1; Page 47; 90pp; English.
 PS The present invention describes a composition (A) comprising an
 CC antimicrobial cationic peptide (I), a viscosity-increasing agent (II) and
 CC a solvent (III). Also described is a composition comprising (I), buffer
 CC (IV) and (III). (I) has antibacterial, virucide, antiinflammatory,
 CC fungicide, protozoacide, parasiticide, vulnery, dermatological,
 CC herbicide and insecticide activities. (A) can be used to reduce the
 CC population of microflora (eukaryotes, prokaryotes or viruses) at a target
 CC site, particularly for treatment or prevention of infections. They can be
 CC used to treat a wide range of systemic infections (e.g. sepsis) and for
 CC topical treatment of wounds, but most especially can be used: (i) at
 CC sites where medical devices have been, or will be, inserted into the body
 CC (alternatively, they are used to treat the devices); and (ii) at sites on
 CC the skin (particularly for treating acne) or the mucosa. The devices
 CC treated are especially central venous, vascular dialysis, pulmonary
 CC artery, peritoneal dialysis or umbilical catheters. They may also be used
 CC as cosmetics and soaps; as herbicides and insecticides; in building
 CC materials (e.g. silicone sealants) and in processing animal products,
 CC e.g. hides. The present sequence represents an antimicrobial cationic
 CC peptide, which is used in the exemplification of the present invention.
 XX Sequence 13 AA;
 SQ

Query Match 100.0%; Score 91; DB 6; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWFWWPRRK 13
 |||||
 DB 1 ILKKWFWWPRRK 13
 |||||

RESULT 15
 ADA00555
 ID ADA00555 standard; peptide; 13 AA.
 XX
 AC ADA00555;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Antimicrobial cationic peptide 11E3CN.
 XX
 KW antimicrobial; cationic; viscosity-increasing agent; solvent; buffer;
 KW antibacterial; virucide; antiinflammatory; fungicide; protozoacide;
 KW parasiticide; vulnery; dermatological; herbicide; insecticide;
 KW infection; systemic infection; sepsis; acne; disinfectant; herbicide;
 KW insecticide; silicone sealant.
 XX
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT Modified-site 13
 FT /label= amidated
 XX
 PN WO2003015809-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 21-AUG-2002; 2002WO-US026525.

XX 21-AUG-2001; 2001US-0314232P.
PR 20-AUG-2002; 2002US-00225087.
XX
XX PA (MICR-) MICROLOGIX BIOTECH INC.
XX Krieger TJ, McNicol PJ, Fraser JR;
XX WPI; 2003-332767/31.
DR
XX Composition containing stabilized antimicrobial cationic protein, useful
PT for treating infections, particularly where associated with in-dwelling
PT devices.
XX
PS Claim 47; Page 48; 90pp; English.
XX
CC The present invention describes a composition (A) comprising an
CC antimicrobial cationic peptide (I), a viscosity-increasing agent (II) and
CC a solvent (III). Also described is a composition comprising (I), buffer
CC (IV) and (III). (I) has antibacterial, virucide, antiinflammatory,
CC fungicide, protozoacide, parasiticide, vulnery, dermatological,
CC herbicide and insecticide activities. (A) can be used to reduce the
CC population of microflora (eukaryotes, prokaryotes or viruses) at a target
CC site, particularly for treatment or prevention of infections. They can be
CC used to treat a wide range of systemic infections (e.g. sepsis) and for
CC topical treatment of wounds, but most especially can be used: (i) at
CC sites where medical devices have been, or will be, inserted into the body
CC (alternatively, they are used to treat the devices); and (ii) at sites on
CC the skin (particularly for treating acne) or the mucosa. The devices
CC treated are especially central venous, vascular dialysis, pulmonary
CC artery, peritoneal dialysis or umbilical catheters. They may also be used
CC as surface disinfectants; for treatment of clothing and air filters; in
CC cosmetics and soaps; as herbicides and insecticides; in building
CC materials (e.g. silicone sealants) and in processing animal products,
CC e.g. hides. The present sequence represents an antimicrobial cationic
CC peptide, which is used in the exemplification of the present invention.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 91; DB 6; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKQWVWVPRK 13
| | | | | | | | | | | | |
DB 1 ILKKQWVWVPRK 13
| | | | | | | | | | | | |

Search completed: May 4, 2004, 15:19:39
Job time : 50.6053 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:14:57 ; Search time 34.5526 seconds
(without alignments)
118.710 Million cell updates/sec

Title: US-09-444-281-85
Perfect score: 99
Sequence: 1 ILPWKPWWPWR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	64	64.6	83 11 Q80VT9	Q80VT9 mus musculus
2	59	59.6	780 16 Q8PE33	Q8PE33 xanthomonas
3	57	57.6	723 12 Q9DUC4	Q9DUC4 tt virus. o
4	55	55.6	49 12 Q9DT80	Q9DT80 tt virus. o
5	55	55.6	152 2 Q8RPF4	Q8RPF4 desulfotoba
6	55	55.6	208 16 Q8PB17	Q8PB17 xanthomonas
7	55	55.6	342 4 Q86BE4	Q86BE4 homo sapien
8	55	55.6	748 12 Q9DTE1	Q9DTE1 tt virus. o
9	55	55.6	750 12 Q9ID04	Q9ID04 tt virus. o
10	54	54.5	1383 12 Q84712	Q84712 porcine epi
11	54	54.5	1383 12 Q91AV1	Q91AV1 porcine epi
12	54	54.5	1383 12 Q8B482	Q8B482 porcine epi
13	54	54.5	1386 12 Q8QQ98	Q8QQ98 porcine epi
14	53.5	54.0	299 4 Q94N1	Q94N1 homo sapien
15	53	53.5	216 5 Q9476	Q9476 drosophila
16	53	53.5	257 17 Q8TW99	Q8TW99 methanopyru

17	53	53.5	328 12 Q8B9N6	Q8B9N6 rachiplusia
18	53	53.5	331 12 Q92380	Q92380 bombyx mori
19	53	53.5	600 5 Q8IGB8	Q8IGB8 drosophila
20	53	53.5	1245 3 Q9Y7V5	Q9Y7V5 trichoderma
21	52.5	53.0	640 2 Q934J3	Q934J3 prevotella
22	52	52.5	102 16 Q8P4Z9	Q8P4Z9 xanthomonas
23	52	52.5	105 16 Q8PPU5	Q8PPU5 xanthomonas
24	52	52.5	282 16 Q8HH3	Q8HH3 pseudomonas
25	52	52.5	351 16 Q8DJH5	Q8DJH5 synechococ
26	52	52.5	388 16 Q7VI01	Q7VI01 helicobacte
27	51.5	52.0	350 16 Q82HM2	Q82HM2 streptomyce
28	51	51.5	55 8 Q9B6T0	Q9B6T0 eudromia el
29	51	51.5	298 17 Q8ZU59	Q8ZU59 pyrobaculum
30	51	51.5	403 16 Q7V805	Q7V805 prochloroco
31	51	51.5	530 4 Q13161	Q13161 homo sapien
32	51	51.5	689 16 Q8YV85	Q8YV85 anabaena sp
33	50.5	51.0	214 5 Q9N9T4	Q9N9T4 leishmania
34	50.5	51.0	970 12 Q9YW19	Q9YW19 melanoplus
35	50.5	51.0	988 12 Q9LHP7	Q9LHP7 ocelluleus as
36	50	50.5	55 8 Q8SEB4	Q8SEB4 elaeinia fal
37	50	50.5	83 16 Q9WYF1	Q9WYF1 thermotoga
38	50	50.5	137 10 Q84ST7	Q84ST7 oryza sativ
39	50	50.5	327 10 Q9AUN3	Q9AUN3 oryza sativ
40	50	50.5	327 10 Q7XFD1	Q7XFD1 oryza sativ
41	50	50.5	337 16 Q92VQ2	Q92VQ2 rhizobium m
42	50	50.5	466 4 Q75035	Q75035 homo sapien
43	49.5	50.0	746 12 Q9JH31	Q9JH31 tt virus. o
44	49.5	50.0	157 5 Q9YOE8	Q9YOE8 drosophila
45	49.5	50.0	198 10 Q8GZX7	Q8GZX7 oryza sativ

ALIGNMENTS

RESULT 1

Q80VT9 PRELIMINARY; PRT; 83 AA.

AC Q80VT9; 01-JUN-2003 (TREMREL. 24, Created)
 DT 01-JUN-2003 (TREMREL. 24, Last sequence update)
 DT 01-JUN-2003 (TREMREL. 24, Last annotation update)
 DE Krtap16.10 (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB; TISSUE=Scapular skin;
 RX MEDLINE=21185977; PubMed=11290294;
 RA Tkatchenko A.V., Visconti R.P., Shang L., Papenbrock T., Pruett N.D.,
 RA Ito T., Ogawa M., Awgulewitsch A.,
 RT "Overexpression of Hoxc13 in differentiating keratinocytes results in
 RL downregulation of a novel hair keratin gene cluster and alopecia.";
 RL Development 128:1547-1558(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB; TISSUE=Scapular skin;
 RA Tkatchenko A.V., Pruett N.D., Awgulewitsch A.,
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF477980; AAC3570.1.1. -
 FT NON TER 1
 SQ SEQUENCE 83 AA; 10834 MW; F7BB37E12A327BCD CRC64;

Query Match 64.6%; Score 64; DB 11; Length 83;
 Best Local Similarity 77.8%; Pred. No. 0.6;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 PWKPWWPW 11

Db 23 PNLWPPWLW 31

```

RESULT 2
QSPF93
ID QSPF93 PRELIMINARY; PRT; 780 AA.
AC QSPF93;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein XCC0088.
GN XCC0088.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2202145; PubMed=1202417;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Seta J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezsa R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL, AF012102; AM39407.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 780 AA; 85074 MW; 12867434D1852549 CRC64;

Query Match 59.6%; Score 59; DB 16; Length 780;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 WKPFWFW 11
DB 148 WPKFWFW 155

RESULT 3
Q9DUC4 PRELIMINARY; PRT; 723 AA.
AC Q9DUC4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF1.
OS TT virus.
OC Viruses; ssDNA viruses; Circoviridae; Anellovirus.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mf-TTV9;
RA Okamoto H.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Mf-TTV9;
RX MEDLINE=20534983; PubMed=11080484;
RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,
RA Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;
RT "Species-specific TT viruses in humans and nonhuman primates and their
RT phylogenetic relatedness.";
RL Virology 277:368-378(2000).

DR EMBL, AB041959; BAB19313.1; --.
DR GO; GO:0004185; P:serine carboxypeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001563; Peptidase_S10.
DR Pfam; PF02956; TT_Virus_Unk.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
SQ SEQUENCE 723 AA; 85393 MW; 232D003098766344 CRC64;

Query Match 57.6%; Score 57; DB 12; Length 723;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PWPFWFR 13
DB 2 PWPFWFR 8

RESULT 4
Q9DT80 PRELIMINARY; PRT; 49 AA.
AC Q9DT80;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE ORF1 (Fragment).
OS TT virus.
OC Viruses; ssDNA viruses; Circoviridae; Anellovirus.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TYM9;
RX MEDLINE=20568739; PubMed=11118348;
RA Okamoto H., Nishizawa T., Tawara A., Takahashi M., Kishimoto J.,
RA Sai T., Sugai Y.;
RT "TT virus mRNAs detected in the bone marrow cells from an infected
RT individual.";
RL Biochem. Biophys. Res. Commun. 279:700-707(2000).
DR EMBL; AB050449; BAB1930.1; --.
FT NON TER 49
SQ SEQUENCE 49 AA; 7225 MW; 1DA6F8F1AB69AA43 CRC64;

Query Match 55.6%; Score 55; DB 12; Length 49;
Best Local Similarity 44.4%; Pred. No. 4.4;
Matches 8; Conservative 1; Mismatches 3; Indels 6; Gaps 1;

QY 2 LPWKWFW -----WFWFR 13
DB 1 MAWTWFWQRRRRWFWFR 18

RESULT 5
Q8RPF4 PRELIMINARY; PRT; 152 AA.
AC Q8RPF4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Desulfotobacterium hafnense.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Peptococcaceae;
OC Desulfotobacterium.
OX NCBI_TaxID=49338;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DCB-2;
RA Davis J.K., Tiedje J.M.;
RT "Sequence and transcriptional analysis of reductive dehalogenase genes
RT of Desulfotobacterium.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF403185; AAL87800.1; --.
DR InterPro; IPR006311; Tat.
DR TIGRPFAMs; TIGR01409; TAT_signal_seq; 1.

```

```

KW Hypothetical protein.
SQ SEQUENCE 152 AA; 16876 MW; 2F5A00F01E70A379 CRC64;

Query Match 55.6%; Score 55; DB 2; Length 152;
Best Local Similarity 85.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 PWKWPWW 9
DB 146 PWIWPWW 152

RESULT 6
Q9PBI7 PRELIMINARY; PRT; 208 AA.
AC Q9PBI7;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein XCC1132.
GN XCC1132.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=3340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=20202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Fatah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
RA Camarotte G., Cannavan F., Cardoso J., Chamberg F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorfi H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Teai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:453-463(2002).
DR EMBL; AE012212; AAM40431.1; -.
DR GO; GO:0009002; F:serine-type D-Ala-D-Ala carboxypeptidase ac. .; IEA.
DR InterPro; IPR003709; Pept_M15B_M15C.
DR Pfam; PF02557; Vany; 1.
KW Hypothetical protein, Complete proteome.
SQ SEQUENCE 208 AA; 22940 MW; 1D5180F6EAF7B014 CRC64;

Query Match 55.6%; Score 55; DB 16; Length 208;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 PWKWPWW 10
DB 200 PWHWRWWP 207

RESULT 7
Q96BE4 PRELIMINARY; PRT; 342 AA.
AC Q96BE4;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin, and amelanotic;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
KW EMBL; BC015687; AAI15687.1; -.
KW Hypothetical protein.
SQ SEQUENCE 342 AA; 37741 MW; 3147596F8D7DF849 CRC64;

Query Match 55.6%; Score 55; DB 4; Length 342;
Best Local Similarity 63.8%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILPWKPWWPW 11
DB 296 LIFGPWWPW 306

RESULT 8
Q9DT81 PRELIMINARY; PRT; 748 AA.
AC Q9DT81;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE ORF1.
OS TT virus.
OC Viruses; ssDNA viruses; Circoviridae; Anellovirus.
OX NCBI_TaxID=68897;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TYM9;
RX MEDLINE=20568739; PubMed=1118348;
RA Okamoto H., Nishizawa T., Tawara A., Takahashi M., Kishimoto J.,
RA Sai T., Sugai Y.;
RT "TT virus RNAs detected in the bone marrow cells from an infected
RT individual.";
RL Biochem. Biophys. Res. Commun. 279:700-707(2000).
DR EMBL; AB050448; BAB19928.1; -.
DR InterPro; IPR004219; TTvirus Unk.
DR Pfam; PF02956; TT_ORF1; 1.
SQ SEQUENCE 748 AA; 88552 MW; D65CCB2CAA5CE26F CRC64;

Query Match 55.6%; Score 55; DB 12; Length 748;
Best Local Similarity 44.4%; Pred. No. 46;
Matches 8; Conservative 1; Mismatches 3; Indels 6; Gaps 1;

QY 2 LPWKWPW-----WPWRR 13
DB 1 MAWTWWQRRRRRWPPWR 18

RESULT 9
Q91D04 PRELIMINARY; PRT; 750 AA.
AC Q91D04;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE ORF1.
OS TT virus.
OC Viruses; ssDNA viruses; Circoviridae; Anellovirus.
OX NCBI_TaxID=68897;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21488921; PubMed=11601907;
RA Okamoto H., Nishizawa T., Takahashi M., Asabe S., Tsuda F.,
RA Yoshikawa A.;
RT "Heterogeneous distribution of TT virus of distinct genotypes in
RT multiple tissues from infected humans.";
RL Virology 288:358-368(2001).

```



```

RA Yeo S.-G., Krell P., Nagy E.;
RT "Cloning and nucleotide sequence analysis of spike gene of porcine
RT epidemic diarrhoea virus detected in Korea.";
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY167585; AAN86621.1; -.
DR InterPro; IPR002551; Corona_S1.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01600; Corona_S1; 1.
DR Pfam; PF01601; Corona_S2; 1.
SQ SEQUENCE 1383 AA; 151582 MW; B5BA4D7EE5371A54 CRC64;

Query Match          54.5%; Score 54; DB 12; Length 1383;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KWPWWPW 11
DB 1322 KWPWWPW 1328

RESULT 13
Q8QQ98
ID Q8QQ98 PRELIMINARY; PRT; 1386 AA.
AC Q8QQ98;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Spike protein.
GN SPK1.
OS Porcine epidemic diarrhoea virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=28295;
RN [1]_TaxID=28295;
RP SEQUENCE FROM N.A.
RA Kang T.-J., Lim Y.-Y., Jang Y.-S., Kwon T.-H., Kim D.-H., Yang M.-S.;
RA "Spike Protein gene of Korea Porcine Epidemic Diarrhoea Virus.";
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF500215; AAM19716.1; -.
DR InterPro; IPR002551; Corona_S1.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01600; Corona_S1; 1.
DR Pfam; PF01601; Corona_S2; 1.
SQ SEQUENCE 1386 AA; 151853 MW; 11F98BCB2AA0526B CRC64;

Query Match          54.5%; Score 54; DB 12; Length 1386;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KWPWWPW 11
DB 1325 KWPWWPW 1331

RESULT 14
Q9YAN1
ID Q9YAN1 PRELIMINARY; PRT; 299 AA.
AC Q9YAN1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKF2P434C192.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RA Ansong W., Wirkner U., Mewes H.W., Gassenhuber J., Wiemann S.;
RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL096753; CAB46428.2; -.

```

```

DR PIR; T12505; T12505.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 299 AA; 34032 MW; 6B8DB60E6A88239A CRC64;

Query Match          54.0%; Score 53.5; DB 4; Length 299;
Best Local Similarity 57.1%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 3 PW---KWPWWPWR 13
DB 30 PWTGSASPWPMSR 43

RESULT 15
Q9W476
ID Q9W476 PRELIMINARY; PRT; 216 AA.
AC Q9W476;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG15768 protein.
GN CG15768.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Berkeley;
RA MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George B.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadien E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Durbin K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-X., Wasarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RT Science 287:2185-2195(2000).
DR EMBL; AE003435; AAF46082.1; -.
DR FlyBase; FBgn0029806; CG15768.
SQ SEQUENCE 216 AA; 24969 MW; B96B4247F9B2924CB CRC64;

```


Query Match 53.5%; Score 53; DB 5; Length 216;
Best Local Similarity 77.8%; Pred. NO. 28;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 LPWKWFWWP 10
| | | | |
Db 62 LPWKWAWLP 70

Search completed: May 4, 2004, 15:22:11
Job time : 35.886 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:15:37 ; Search time 11.9737 Seconds
(without alignments)
104.437 Million cell updates/sec

Title: US-09-444-281-35

Perfect score: 91

Sequence: 1 ILKKPWPWPWRK 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	80.2	144	1 JCI222	indolicidin precu
2	54	59.3	1173	1 VGIHHC	E2 glycoprotein pr
3	51	56.0	299	2 T12505	hypothetical prote
4	49	53.8	298	2 B72492	hypothetical prote
5	49	53.8	467	2 B89605	protein F18G5.2 [i
6	49	53.8	498	1 J70751	ferridoxin-NADP re
7	49	53.8	527	2 S33088	myosin heavy chain
8	49	53.8	715	2 B70741	probable moey prot
9	49	53.8	1940	2 A59287	myosin heavy chain
10	48	52.7	111	2 T29295	hypothetical prote
11	48	52.7	265	2 AH0755	conserved hypotet
12	47.5	52.2	114	2 T36208	hypothetical prote
13	47	51.6	248	2 S23449	NADH oxidase [H2O2
14	47	51.6	253	2 G70715	hypothetical prote
15	46.5	51.1	352	2 S77354	histidinol-phospha
16	46.5	51.1	621	2 S37664	peplomeric polypro
17	46.5	51.1	630	2 S37663	peplomeric polypro
18	46.5	51.1	1154	1 VGIHIB	E2 glycoprotein pr
19	46.5	51.1	1162	1 VGIHAK	E2 glycoprotein pr
20	46.5	51.1	1162	2 S07421	E2 glycoprotein pr
21	46.5	51.1	1162	2 S14939	E2 glycoprotein pr
22	46.5	51.1	1162	2 S14940	E2 glycoprotein pr
23	46	50.5	196	2 S55483	modulator of drug
24	46	50.5	617	2 T22175	hypothetical prote
25	46	50.5	623	2 T22177	hypothetical prote
26	46	50.5	1333	2 S55812	RNA-directed DNA p
27	45	49.5	67	2 AC1954	hypothetical bio
28	45	49.5	273	2 F82646	monofunctional bio
29	45	49.5	276	2 B83161	monofunctional bio

ALIGNMENTS

RESULT 1

JCI222

indolicidin precursor - bovine

N;Alternate names: antimicrobial peptide

C;Species: Bos primigenius taurus (cattle)

C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999

C;Accession: JCI222; A42387; S25664

Ridel Sal, G.; Storici, P.; Schneider, C.; Romeo, D.; Zanetti, M.

Biochem. Biophys. Res. Commun. 187, 467-472, 1992

A;Title: CDNA cloning of the neutrophil bactericidal peptide indolicidin.

A;Reference number: JCI222; MUID:92392368; PMID:1520337

A;Accession: JCI222

A;Molecule type: mRNA

A;Residues: 1-144 <SAL>

A;Cross-references: EMBL:X67340; NID:G462; PIDN:CAA47755.1; PID:G463

A;Experimental source: bone marrow

R;Selected, M.E.; Novotny, M.J.; Morris, W.L.; Tang, Y.Q.; Smith, W.; Cullor, J.S.

J. Biol. Chem. 267, 4292-4295, 1992

A;Title: Indolicidin, a novel bactericidal tridecapeptide amide from neutrophils.

A;Reference number: A42387; MUID:92165771; PMID:1537821

A;Accession: A42387

A;Molecule type: protein

A;Residues: 131-143 <SEL>

A;Experimental source: neutrophils

A;Note: sequence extracted from NCBI backbone (NCBIP:83840)

C;Superfamily: cathelin; cystatin homology

C;Keywords: amidated carboxyl end

F;1-29/Domain: signal sequence #status predicted <SIG>

F;22-129/Domain: cystatin homology <CYS>

F;30-130/Domain: propeptide #status predicted <PRO>

F;131-143/Product: indolicidin #status experimental <MAT>

F;143/Modified site: amidated carboxyl end (Arg) (amide in mature form from following g

Query Match 80.2%; Score 73; DB 1; Length 144;

Best Local Similarity 100.0%; Pred. No. 0.0034;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KPWPFWR 12

DB 135 KPWPFWR 143

RESULT 2

VGIHHC

E2 glycoprotein precursor - human coronavirus (strain 229E)

N;Alternate names: peplomeric glycoprotein; S glycoprotein; spike glycoprotein

C;Species: human coronavirus

A;Note: host Homo sapiens (man)

C;Date: 31-Dec-1991 #sequence revision 31-Dec-1991 #text_change 16-Jun-2000

C;Accession: A34766; S05460

R;Raabe, T.; Schelle-Prinz, B.; Siddell, S.G.

J. Gen. Virol. 71, 1065-1073, 1990

A;Title: Nucleotide sequence of the gene encoding the spike glycoprotein of human corona
A;Reference number: A34766; MUID:90264837; PMID:2345367
A;Accession: A34766
A;Molecule type: mRNA
A;Residues: 1-1173 <RAA>
A;Cross-references: EMBL:X16816; NID:g58926; PIDN:CAA34723.1; PID:g58927
A;Experimental source: strain 229E
R;Raabe, T.; Siddell, S.
Nucleic Acids Res 17, 6387, 1989
A;Title: Nucleotide sequence of the human coronavirus HCV 229E mRNA 4 and mRNA 5 unique
A;Reference number: A34038; MUID:89366667; PMID:2701946
A;Accession: S05460
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1159-1173 <RA2>
A;Cross-references: EMBL:X15654; NID:g58921; PIDN:CAA33680.1; PID:gl334827
C;Superfamily: coronavirus E2 glycoprotein
C;Keywords: glycoprotein; transmembrane protein
F;15/Domain: signal sequence #status predicted <SIG>
F;16-1173/Product: E2 glycoprotein #status predicted <MAT>
F;116-1138/Domain: transmembrane #status predicted <TMN>
F;23,62,98,147,171,176,220,243,326,333,440,464,518,538,542,568,581,587,663,671,930,1015,
1173
Query Match 59.3%; Score 54; DB 1; Length 1173;
Best Local Similarity 85.7%; Pred. No. 9.1;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 KWPWPWP 10
|||||
Db 1113 KWPWPWP 1119
RESULT 3
T12505
hypothetical protein DKFPz434C192.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 23-Jul-1999
C;Accession: T12505
R;Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z17527
A;Accession: T12505
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-299 <ANS>
A;Cross-references: EMBL:AL096753
A;Experimental source: adult testis; clone DKFPz434C192
C;Genetics:
A;Note: DKFPz434C192.1
Query Match 56.0%; Score 51; DB 2; Length 299;
Best Local Similarity 85.7%; Pred. No. 6.1;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6 PWPWPWR 12
|||||
Db 37 PWPWPWR 43
RESULT 4
B72492
hypothetical protein APB2577 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: B72492
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yanazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yanazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: B72492
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-298 <RAW>
A;Cross-references: DDBJ:AF000064; NID:g5105945; PIDN:BAA81594.1; PID:d1045380; PID:g51
A;Experimental source: strain K1
C;Genetics:
A;Gene: APB2577
Query Match 53.8%; Score 49; DB 2; Length 298;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 LKKWPWPWP 11
|||||
Db 102 IKETPWPWR 111
RESULT 5
E89605
Protein F18G5.2 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Nov-2001
C;Accession: E89605
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_el
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; ar
A;Accession: E89605
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-467 <STO>
A;Cross-references: GB:chr_X; PIDN:AAA81082.1; PID:g1055093; GSPDB:GN00028; CESP:F18G5.
C;Genetics:
A;Map position: X
Query Match 53.8%; Score 49; DB 2; Length 467;
Best Local Similarity 83.3%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 WFPWPWP 10
|||||
Db 201 WFPWPWP 206
RESULT 6
JT0751
ferredoxin-NADP reductase (EC 1.18.1.2), long form precursor - bovine
N;Alternate names: adrenodoxin reductase
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Jul-1994 #sequence_revision 18-Oct-1996 #text_change 03-Jun-2002
C;Accession: JT0751; JT0079; J30390; S03558; P50003; A29604; S52100
R;Takata, Y.; Sagara, Y.; Kono, A.; Sekimizu, K.; Horiuchi, T.
Biol. Pharm. Bull. 16, 1200-1206, 1993
A;Title: Gene structure of bovine adrenodoxin reductase.
A;Reference number: JT0751; MUID:94177140; PMID:8130767
A;Accession: JT0751
A;Molecule type: DNA
A;Residues: 1-498 <TAK>
A;Cross-references: GB:D83475; NID:g1199916; PIDN:BAA11921.1; PID:g4521308
A;Experimental source: adrenal cortex
A;Note: the authors translated the codon GTC for residue 205 as Gly
R;Sagara, Y.; Takata, Y.; Miyata, T.; Hara, T.; Horiuchi, T.
J. Biochem. 102, 1333-1336, 1987
A;Title: Cloning and sequence analysis of adrenodoxin reductase cDNA from bovine adrena
A;Reference number: JT0079; MUID:88198050; PMID:3448086
A;Accession: JT0079
A;Molecule type: mRNA
A;Residues: 1-204,211-498 <SAG>
A;Cross-references: GB:D00211; NID:g217433; PIDN:BAA00150.1; PID:g217434
A;Note: the deduced sequence is partially confirmed by amino acid sequencing of 15 isol
R;Sagara, Y.
submitted to DDBJ, September 1989

A;Reference number: J50390
 A;Contents: revision, insertion of residues 205-210
 A;Accession: J50390
 A;Molecule type: mRNA
 A;Residues: 56-498 <SA2>
 R;Hanukoglu, I.; Gutfinger, T.
 Eur. J. Biochem. 180, 479-484, 1989
 A;Title: cDNA sequence of adrenodoxin reductase. Identification of NADP-binding sites in
 A;Reference number: S03558; MUID:89170752; PMID:2924777
 A;Accession: S03558
 A;Molecule type: mRNA
 A;Residues: 153-204,211-498 <HAN>
 A;Cross-references: ENBL:X13736; NID:G65; PIDN:CAA32002.1; PID:G933776
 A;Note: 405-Ser was also found
 R;Hamamoto, I.; Kurokouchi, K.; Tanaka, S.; Ichikawa, Y.
 Biochim. Biophys. Acta 953, 207-213, 1988
 A;Title: Adrenodoxin-binding peptide of NADPH-adrenodoxin reductase.
 A;Reference number: P50003; MUID:88184054; PMID:3355838
 A;Accession: P50003
 A;Molecule type: protein
 A;Residues: 33-41,'S',43-62;260-283,'TM',496-498 <HAM>
 A;Note: a cyanogen bromide peptide binds to adrenodoxin
 R;Nonaka, Y.; Murakami, H.; Yabasaki, Y.; Kuramitsu, S.; Kagamiyama, H.; Yamano, T.; Oka
 Biochem. Biophys. Res. Commun. 145, 1239-1247, 1987
 A;Title: Molecular cloning and sequence analysis of full-length cDNA for mRNA of adrenod
 A;Reference number: A29604; MUID:87270696; PMID:3038094
 A;Accession: A29604
 A;Molecule type: mRNA
 A;Residues: 1-76,'R',78-80,'VWLALTTTPSRML',95-123,'RVVRLT',129-204,211-273,'R',275-322,
 A;Cross-references: GB:M7029; NID:G162628; PIDN:AAA30362.1; PID:G162629
 A;Experimental source: adrenal cortex
 R;Warburton, R.J.; Seybert, D.W.
 Biochim. Biophys. Acta 1246, 39-46, 1995
 A;Title: Structural and functional characterization of bovine adrenodoxin reductase by 1
 A;Reference number: S52100; MUID:95110846; PMID:7811729
 A;Accession: S52100
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 'X',34-41,'X',43-48,'X',50-51;304-306,'X',308-309,'X',311-326 <WAR>
 A;Comment: Ferredoxin-NADP+ reductase is localized in the matrix of adrenal cortex mito
 erredoxin-NADP+ reductase, adrenodoxin and two forms of cytochrome P-450.
 C;Genetics:
 A;Introns: 27/1; 59/3; 91/3; 132/3; 170/3; 204/3; 246/3; 275/1; 341/3; 399/1; 456/1
 C;Function:
 A;Description: catalyzes the reversible reduction of NADP+ by reduced ferredoxin or redu
 C;Superfamily: human ferredoxin-NADP+ reductase
 C;Keywords: alternative splicing; flavoprotein; mitochondrion; monomer; NADP; oxidoreduc
 F1-32/Domain: transit peptide (mitochondrion) #status predicted <SIG>
 F33-498/Product: ferredoxin-NADP+ reductase, long form #status predicted <MAT>
 F33-204,211-498/Product: ferredoxin-NADP+ reductase, short form #status experimental <M
 F40-70/Region: beta-alpha-beta FAD nucleotide-binding fold
 F180-190/Region: NADP binding #status predicted
 F281/Binding site: substrate (lys) #status experimental
 Query Match 53.8%; Score 49; DB 1; Length 498;
 Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 5 WPNWPNW 10
 DB 6 WRWPNW 11
 RESULT 7
 S33068
 myosin heavy chain - fluke (Schistosoma mansoni) (fragment)
 N;Alternate names: surface antigen, 200K
 C;Species: Schistosoma mansoni
 C;Date: 22-Nov-1993 #sequence_revision 06-Sep-1996 #text_change 13-Feb-1998
 C;Accession: S33068
 R;Soisson, L.M.A.; Masterson, C.P.; Tom, T.D.; McNally, M.T.; Lowell, G.H.; Strand, M.
 J. Immunol. 149, 3612-3620, 1992
 A;Title: Induction of protective immunity in mice using a 62-kDa recombinant fragment of

A;Reference number: A46514; MUID:93056536; PMID:1431131
 A;Accession: S33068
 A;Molecule type: mRNA
 A;Residues: 1-527 <SOI>
 A;Cross-references: ENBL:X65591
 A;Note: the authors translated the codon CAA for residue 346 as Lys
 C;Superfamily: myosin heavy chain; myosin motor domain homology
 C;Keywords: ATP; surface antigen

Query Match 53.8%; Score 49; DB 2; Length 527;
 Best Local Similarity 62.5%; Pred. No. 20;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ILKKWPNW 8
 DB 106 VLKNWPNW 113

RESULT 8

B70741
 probable moey protein - Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C;Accession: B70741
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Nature 393, 537-544, 1998
 A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: B70741
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-715 <COL>
 A;Cross-references: GB:Z75555; GB:AL123456; NID:G3261608; PIDN:CAA99988.1; PID:62503556;
 A;Experimental source: strain H37RV
 C;Genetics:
 A;Gene: moey

Query Match 53.8%; Score 49; DB 2; Length 715;
 Best Local Similarity 60.0%; Pred. No. 26;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 KKWPNWPNWR 12
 DB 64 KRWAYTPWR 73

RESULT 9

A59287
 myosin heavy chain - fluke (Schistosoma mansoni) (strain Brazilian LE)
 C;Species: Schistosoma mansoni
 C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 08-Sep-2000
 C;Accession: A59287
 R;Weston, D.S.; Schmitz, J.; Kemp, M.; Kunz, W.
 Mol. Biochem. Parasitol. 58, 161-164, 1993
 A;Title: Cloning and sequence characterization of a complete myosin heavy chain cDNA fr
 A;Reference number: A59287; MUID:93211444; PMID:8459827
 A;Accession: A59287
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-1940 <WES>
 A;Cross-references: GB:I01634; PIDN:AAA29905.1
 A;Experimental source: strain Brazilian LE
 C;Genetics:
 A;Gene: MYH
 C;Superfamily: myosin heavy chain; myosin motor domain homology
 F;82-752/Domain: myosin motor domain homology <WMO>

Query Match 53.8%; Score 49; DB 2; Length 1940;
 Best Local Similarity 62.5%; Pred. No. 69;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILKQWFW 8
: : : : :
DB 809 VLKQWFW 816

RESULT 10
T23295
hypothetical protein C50F7.8 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T23295
R/Johnson, D.; Stellyes, L.
A/Description: The sequence of C. elegans cosmid C50F7.
A/Reference number: Z20601
A/Accession: T23295
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-111 <JOH>
A/Cross-references: EMBL:U41557; PIDN:AAA83303.1; CBSP:C50F7.8
C/Genetics:
A/Gene: CBSP:C50F7.8

Query Match 52.7%; Score 48; DB 2; Length 111;
Best Local Similarity 75.0%; Pred. No. 6;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 WPKWFWRR 12
: : : : :
DB 15 WPKWFWRR 22

RESULT 11
AH0755
conserved hypothetical protein STY2208 [imported] - Salmonella enterica subsp. enterica
C/Species: Salmonella enterica subsp. enterica serovar Typhi
A/Note: this species has also been called Salmonella typhi
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: AH0755
R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moulle, S.; O'Garra, P.
Nature 413, 848-852, 2001
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A/Reference number: AB0502; MUID:21534947; PMID:11677608
A/Accession: AH0755
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-265 <PAR>
A/Cross-references: GB:AL513382; PIDN:CAD05747.1; PID:g16503239; GSPDB:GN00176
C/Genetics:
A/Gene: STY2208

Query Match 52.7%; Score 48; DB 2; Length 265;
Best Local Similarity 41.2%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 4 KWPW-----WFW 10
: : : : :
DB 3 KWPWKAQETQNEFW 19

RESULT 12
T36208
hypothetical protein SC36.09 - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C/Accession: T36208
R/Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A/Reference number: Z21601

A/Accession: T36208
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-114 <OLI>
A/Cross-references: EMBL:AL049763; PIDN:CA842078.1; GSPDB:GN00070; SCOEDB:SC36.09
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: SCOEDB:SC36.09

Query Match 52.2%; Score 47.5; DB 2; Length 114;
Best Local Similarity 63.6%; Pred. No. 7.1;
Matches 7; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 3 KKW-PWMPWR 12
: : : : :
DB 102 RWRPRPWR 112

RESULT 13
S23449
NADH oxidase (H2O2-forming) (EC 1.6.-.-) - Thermus aquaticus
C/Species: Thermus aquaticus
C/Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 30-Sep-2002
C/Accession: S23449; S24556
R/Park, H.J.; Kreutzer, R.; Reiser, C.O.A.; Sprinzl, M.
Eur. J. Biochem. 205, 875-879, 1992
A/Title: Molecular cloning and nucleotide sequence of the gene encoding a H(2)O(2)-form
A/Reference number: S23449; MUID:92249331; PMID:1577004
A/Accession: S23449
A/Molecule type: DNA
A/Residues: 1-248 <PAR>
A/Cross-references: EMBL:X60110
A/Accession: S24556
A/Molecule type: protein
A/Residues: 1-32 <PAR1>
C/Genetics:
A/Gene: nox
C/Superfamily: NADPH-flavin oxidoreductase homolog
C/Keywords: NAD; oxidoreductase
F:1-248/Product: NADH oxidase (H2O2-forming) #status experimental <MAT>

Query Match 51.6%; Score 47; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PWMPW 10
: : : : :
DB 179 PWMPW 183

RESULT 14
G70715
hypothetical protein Rv0945 - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C/Accession: G70715
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: G70715
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-253 <COL>
A/Cross-references: GB:Z79700; GB:AL123456; NID:g3261628; PIDN:CA802005.1; PID:g1524217
A/Experimental source: strain H37RV
C/Genetics:
A/Gene: Rv0945
C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
F:8-190/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 51.6%; Score 47; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PWPWP 10
|||
Db 230 PWPWP 234

RESULT 15
S77354
histidinol-phosphate aminotransferase hisC-1 - *Synechocystis* sp. (strain PCC 6803)
N/Alternate names: protein ell1713
C/Species: *Synechocystis* sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C/Accession: S77354
R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A/Reference number: S74322; MUID:97061201; PMID:8905231
A/Accession: S77354
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-352 <KAN>
A/Cross-references: EMBL:D90906; GB:AB001339; NID:91652492; PIDN:BAAL7457.1; PID:G165253
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C/Genetics:
A/Gene: hisC-1
C/Superfamily: probable histidinol-phosphate transaminase

Query Match 51.1%; Score 46.5; DB 2; Length 352;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

OY 2 LKKWFW-----WPWRR 12
|||
Db 106 LKTIQWQVDQWPWPK 121

Search completed: May 4, 2004, 15:22:58
Job time : 12.9737 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:08:51 ; Search time 8.21053 seconds
(without alignments)

82.444 Million cell updates/sec

Title: US-09-444-281-35

Perfect score: 91

Sequence: 1 ILKKWPWPWRRK 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	80.2	144	1 INDC BOVIN	P33046 bos taurus
2	54	59.3	1173	1 VGL2_CVH22	P15423 human coron
3	49	53.8	492	1 ADRO_BOVIN	P08165 bos taurus
4	49	53.8	715	1 YD55_MYCTU	Q11025 mycobacteri
5	47	51.6	253	1 Y945_MYCTU	P15844 mycobacteri
6	46.5	51.1	1154	1 VGL2_IBVD2	P12722 avian infec
7	46.5	51.1	1162	1 VGL2_IBVB	P11223 avian infec
8	46.5	51.1	1162	1 VGL2_IBVK	P12650 avian infec
9	46.5	51.1	1162	1 VGL2_IBVM	P12651 avian infec
10	46.5	51.1	1163	1 VGL2_IBV6	P251135 avian infec
11	46	50.5	196	1 YAO5_SCHPO	Q09677 schistosach
12	46	50.5	250	1 NPD_METRA	Q8tw90 methanopyru
13	45	49.5	397	1 NML6_MYCTU	Q10773 mycobacteri
14	45	49.5	505	1 TRPE_PSESS	P21689 pseudomonas
15	45	49.5	512	1 PEN2_YEAST	P25621 saccharomyc
16	45	49.5	964	1 NML5_MYCTU	Q53784 mycobacteri
17	45	49.5	967	1 NML4_MYCTU	Q53735 mycobacteri
18	45	49.5	968	1 NML2_MYCTU	Q11171 mycobacteri
19	45	49.5	1108	1 CN3B_RAT	Q63085 rattus norv
20	45	49.5	1225	1 VGL2_CVPR8	P27655 porcine res
21	45	49.5	1225	1 VGL2_CVPRM	P24413 porcine res
22	45	49.5	1235	1 VGL2_CVPMJH	P11225 murine coro
23	45	49.5	1255	1 VGL2_CVHSA	P59594 human coron
24	45	49.5	1324	1 VGL2_CVMA5	P11224 murine coro
25	45	49.5	1353	1 VGL2_CVHOC	P36334 human coron
26	45	49.5	1363	1 VGL2_CVBF	P25190 bovine coro
27	45	49.5	1363	1 VGL2_CVBL9	P25191 bovine coro
28	45	49.5	1363	1 VGL2_CVBLV	P25192 bovine coro
29	45	49.5	1363	1 VGL2_CVBM	P15777 bovine coro
30	45	49.5	1363	1 VGL2_CVBO	P25193 bovine coro
31	45	49.5	1363	1 VGL2_CVBV	P25194 bovine coro
32	45	49.5	1376	1 VGL2_CVM4	P22432 murine coro
33	45	49.5	1376	1 VGL2_CVMJC	Q02385 murine coro

RESULT 1
ID INDC BOVIN STANDARD; PRT; 144 AA.
AC P33046;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Indolicidin precursor.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=92292368; PubMed=1520337;
RA del Sal G., Storici P., Schneider C., Romeo D., Zanetti M.;
RT "cDNA cloning of the neutrophil bactericidal peptide indolicidin.";
RL Biochem. Biophys. Res. Commun. 187:467-472(1992).
RN [2]
RP SEQUENCE OF 131-143.
RC TISSUE=Neutrophils;
RX MEDLINE=92165771; PubMed=1537821;
RA Sellsted M.E., Novotny M.J., Morris W.L., Tang Y.-Q., Smith W.,
Cullor J.S.;
RT "Indolicidin, a novel bactericidal tridecapeptide amide from
neutrophils.";
RL J. Biol. Chem. 267:4292-4295(1992).
CC -|- FUNCTION: Potent microbicidal activity, active against
Staphylococcus aureus and Escherichia coli.
CC -|- TISSUE SPECIFICITY: Large granules of neutrophils.
CC -|- PTM: Elastase might be responsible for its maturation.
CC -|- SIMILARITY: Belongs to the cathelicidin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

34 45 49.5 1447 1 VGL2_CVPPU P07946 porcine tra
35 45 49.5 1447 1 VGL2_CVPR8 P18450 porcine tra
36 45 49.5 1449 1 VGL2_CVPRS P33470 porcine tra
37 45 49.5 1449 1 VGL2_CVPMI P36300 canine ente
38 45 49.5 1451 1 VGL2_CVCAI P10033 feline infe
39 45 49.5 1452 1 VGL2_FIPV Q97mh9 mus musculu
40 45 49.5 2035 1 M18A_MOUSE Q92614 homo sapien
41 45 49.5 2054 1 M18A_HUMAN P08799 dictyosteli
42 45 49.5 2116 1 MY52_DICDI C81ug5 homo sapien
43 45 49.5 2567 1 M18B_HUMAN P14589 plasmodium
44 48.4 53 1 YDH3_PLAFS P75709 escherichia
45 48.4 151 1 YBBJ_ECOLI

ALIGNMENTS

```

FT PROPEP      30 130
FT PEPTIDE    131 143
FT MOD_RES   30 30
FT
FT DISULFID   85 96
FT DISULFID  107 124
FT MOD_RES   143 143
SQ SEQUENCE  144 AA; 16479 MW; E3BLCBBS5C09911 CRC64;

Query Match      80.2%; Score 73; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KPWFWPWR 12
DB 135 KPWFWPWR 143

RESULT 2
VGL2_CVH22 STANDARD; PRT: 1173 AA.
ID VGL2_CVH22
AC F15423; P89342; P89343; P89344; Q66174; Q990M1; Q990M2; Q990M3;
AC Q990M4;
DC 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein).
GN S.
OS Human coronavirus (strain 229E) (HCoV-229E).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11137;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90264837; PubMed=2345367;
RA Raabe T., Schelle-Prinz B., Siddell S.G.;
RT "Nucleotide sequence of the gene encoding the spike glycoprotein of
RT human coronavirus HCV 229E."
RL J. Gen. Virol. 71:1065-1073(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21262210; PubMed=11369870;
RA Thiel V., Herold J., Schelle B., Siddell S.G.;
RT "Infectious RNA transcribed in vitro from a cDNA copy of the human
RT coronavirus genome cloned in vaccinia virus."
RL J. Gen. Virol. 82:1273-1281(2001).
RN [3]
RP SEQUENCE OF 98-1113 FROM N.A., AND VARIANTS.
RC STRAIN=Isolate RW Stock, Isolate P100E, Isolate P11A, and
RC Isolate P11B;
RA Bonavia A., Holmes K.V.;
RT "Viral and cellular changes in a human cell line persistently infected
RT with human coronavirus HCoV-229E."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 98-1113 FROM N.A., AND VARIANTS.
RC STRAIN=Isolate ATCC VR-74, Isolate A162, and Isolate LRI 281;
RX MEDLINE=99086140; PubMed=9870593;
RA Hays J.P., Wynt S.H.;
RT "PCR sequencing of the spike genes of geographically and
RT chronologically distinct human coronaviruses 229E."
RL J. Virol. Methods 75:179-193(1998).
RN [5]
RP SEQUENCE OF 1159-1173 FROM N.A.
RX MEDLINE=89366667; PubMed=2701946;
RA Raabe T., Siddell S.;
RT "Nucleotide sequence of the human coronavirus HCV 229E mRNA 4 and mRNA
RT 5 unique regions."
RL Nucleic Acids Res. 17:6387-6387(1989).
RN [6]
RP INTERACTION WITH ANPEP.
RX MEDLINE=22440020; PubMed=12551991;
RA Bonavia A., Zelus B.D., Wertworth D.E., Talbot P.J., Holmes K.V.;

```

```

RT "Identification of a receptor-binding domain of the spike glycoprotein
RT of human coronavirus HCoV-229E."
RL J. Virol. 77:2530-2538(2003).
RN [7]
RP INTERACTION WITH ANPEP.
RX MEDLINE=22511439; PubMed=12634402;
RA Breslin J.J., Mork I., Smith M.K., Vogel L.K., Hemmila E.M.,
RA Bonavia A., Talbot P.J., Sjoestrom H., Noren O., Holmes K.V.;
RT "Human coronavirus 229E: receptor binding domain and neutralization by
RT soluble receptor at 37 degrees C."
RL J. Virol. 77:4435-4438(2003).
RN [8]
RP REVIEW.
RX MEDLINE=21109095; PubMed=11162792;
RA Gallagher T.M., Buchmeier M.J.;
RT "Coronavirus spike proteins in viral entry and pathogenesis."
RL Virology 279:371-374(2001).
CC -1- FUNCTION: Structural protein that makes spikes at the surface of
CC the virus. Determines enteropathogenicity and virulence of the
CC virus. Initiates infection by specifically recognizing and binding
CC the human aminopeptidase ANPEP receptor. Its association with
CC ANPEP may lead to its conformational change that triggers fusion
CC between viral and host cellular membrane.
CC -1- SUBUNIT: Homotrimer. During virus morphogenesis, it is found in a
CC complex with M and HE proteins (By similarity). Interacts with
CC ANPEP.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DOMAIN: The spike S1 domain displays the specificity for the host
CC receptor.
CC -1- DOMAIN: The leucine zipper-like heptad repeats may mediate the
CC fusion of viral and cellular membranes.
CC -1- POLYMORPHISM: The strong variation between the different
CC strains may affect the virulence of the virus.
CC -1- MISCELLANEOUS: In contrast to serogroup 2, E2 glycoprotein protein
CC from serogroup 1 is not cleaved.
CC -1- SIMILARITY: Contains 1 spike S1 domain.
CC -1- SIMILARITY: Contains 1 spike S2 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X16816; CAA34723.1; -
CC EMBL; AF304460; AAG48592.1; -
CC EMBL; AF344186; AAK32188.1; -
CC EMBL; AF344187; AAK32189.1; -
CC EMBL; AF344188; AAK32190.1; -
CC EMBL; AF344189; AAK32191.1; -
CC EMBL; Y09923; CAA71056.1; -
CC EMBL; Y10051; CAA71146.1; -
CC EMBL; Y10052; CAA71147.1; -
CC EMBL; X15654; CAA33680.1; -
CC PIR; A34766; VGIHHC.
CC InterPro; IPR002551; Corona_S1.
CC InterPro; IPR002552; Corona_S2.
CC Pfam; PF01600; Corona_S1; 1.
CC Pfam; PF01601; Corona_S2; 1.
CC Virulence; Glycoprotein; Envelope protein; Transmembrane; Signal;
CC Coiled coil.
CC
CC SIGNAL 1 15
CC CHAIN 16 1173
CC DOMAIN 16 1115
CC TRANSMEM 1116 1135
CC DOMAIN 1136 1173
CC DOMAIN 32 536
CC DOMAIN 417 547
CC DOMAIN 537 1171
CC DOMAIN 1054 1103
CC DOMAIN 1067 1102
CC
CC E2 GLYCOPROTEIN.
CC EXTRACELLULAR (POTENTIAL).
CC POTENTIAL.
CC CYTOPLASMIC (POTENTIAL).
CC SPIKE S1.
CC INTERACTION WITH ANPEP.
CC SPIKE S2.
CC COILED COIL (POTENTIAL).
CC LEUCINE ZIPPER-LIKE HEPTAD REPEATS.

```


RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
RX MEDLINE=20455764; PubMed=10998235;
RA Ziegler G.A., Schulz G.E.;
RT "Crystal structures of adrenodoxin reductase in complex with NADP+ and
RT NADPH suggesting a mechanism for the electron transfer of an enzyme
RT family.";
RL Biochemistry 39:10986-10995(2000).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH ADRENODOXIN.
RX MEDLINE=21264735; PubMed=11053423;
RA Mueller J.J., Lapkov A., Bourenkov G., Ruckpaul K., Heinemann U.;
RT "Adrenodoxin reductase-adrenodoxin complex structure suggests electron
RT transfer path in steroid biosynthesis.";
RL J. Biol. Chem. 276:2785-2789(2001).
CC -|- FUNCTION: Serves as the first electron transfer protein in all the
CC mitochondrial p450 systems. Including cholesterol side chain
CC cleavage in all steroidogenic tissues, steroid 11-beta
CC hydroxylation in the adrenal cortex, 25-OH-vitamin D3-24
CC hydroxylation in the kidney, and steroid C-27 hydroxylation in the
CC liver.
CC -|- CATALYTIC ACTIVITY: Reduced adrenodoxin + NADP(+) = oxidized
CC adrenodoxin + NADPH.
CC -|- COFACTOR: FAD.
CC -|- PATHWAY: Cholesterol side-chain-cleavage system.
CC -|- SUBUNIT: Monomer.
CC -|- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -|- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=Short;
CC IsoId=P08165-1; Sequence=Displayed;
CC Name=Long;
CC IsoId=P08165-2; Sequence=VSP_003415;
CC Note=Represents 10-20% of all adrenodoxin reductase mRNAs and
CC seems to be inactive;
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D83475; BAA11921.1; -;
DR EMBL; D83472; BAA11921.1; JOINED.
DR EMBL; D83473; BAA11921.1; JOINED.
DR EMBL; D83474; BAA11921.1; JOINED.
DR EMBL; M17029; AAA30362.1; -;
DR EMBL; D00211; BAA00150.1; -;
DR EMBL; X13736; CAA32002.1; -;
DR PIR; J0751; J0751.
DR PDB; 1CJC; 12-APR-99.
DR PDB; 1E1L; 24-SEP-00.
DR PDB; 1E1K; 24-SEP-00.
DR PDB; 1E1N; 24-SEP-00.
DR PDB; 1E1N; 24-SEP-00.
DR PDB; 1E1N; 24-SEP-00.
DR PDB; 1E6E; 01-AUG-03.
DR InterPro; IPR000759; Adrdx_reductase.
DR PRINTS; PR00419; ADXRDXTASE.
KW Electron transport; Cholesterol metabolism; Oxidoreductase;
KW Mitochondrion; FAD; Flavoprotein; NADP; Transit peptide;
KW Alternative splicing; 3D-structure.
FT TRANSIT 1 32 MITOCHONDRION.
FT CHAIN 33 492 NADPH:ADRENODOXIN OXIDOREDUCTASE.
FT VARSPIC 204 204 E -> EVLLCQC (in isoform long).
FT /FTid=VSP_003415.
FT G -> R (IN REF. 3).
FT CONFLICT 77 77 FGVPADHPVKVNI -> VWLALTTPSRMLL (IN REF.
FT CONFLICT 81 94 3).
FT ODAYH -> RVVBLT (IN REF. 3).
FT CONFLICT 124 128 X -> R (IN REF. 3).
FT CONFLICT 268 268 PS -> RL (IN REF. 3).
FT CONFLICT 317 318 RAAGIRLAVTR -> ARRSAAQSPQE (IN REF. 3).
FT CONFLICT 323 333

FT CONFLICT 341 352
FT STRAND 40 44
FT HELIX 48 60
FT STRAND 65 69
FT TURN 77 77
FT HELIX 78 81
FT TURN 82 82
FT TURN 85 86
FT HELIX 88 92
FT HELIX 93 101
FT TURN 102 102
FT TURN 104 105
FT STRAND 106 110
FT STRAND 114 114
FT TURN 115 117
FT STRAND 118 118
FT HELIX 120 126
FT STRAND 129 132
FT STRAND 138 139
FT TURN 145 148
FT TURN 150 151
FT STRAND 152 154
FT STRAND 155 162
FT HELIX 163 164
FT HELIX 166 168
FT TURN 169 170
FT TURN 175 176
FT STRAND 179 183
FT STRAND 187 197
FT HELIX 200 203
FT TURN 204 205
FT HELIX 210 217
FT TURN 218 218
FT STRAND 223 227
FT HELIX 232 234
FT HELIX 239 246
FT TURN 247 247
FT TURN 249 250
FT STRAND 251 254
FT HELIX 257 260
FT TURN 261 262
FT HELIX 263 266
FT TURN 267 269
FT HELIX 272 286
FT HELIX 291 298
FT TURN 299 299
FT STRAND 302 307
FT STRAND 310 317
FT TURN 319 320
FT STRAND 324 335
FT HELIX 338 340
FT STRAND 342 353
FT STRAND 356 359
FT STRAND 363 364
FT TURN 370 371
FT STRAND 375 375
FT TURN 376 379
FT STRAND 380 380
FT STRAND 383 384
FT TURN 385 386
FT STRAND 387 388
FT TURN 389 390
FT TURN 392 393
FT STRAND 394 396
FT TURN 398 398
FT HELIX 399 402
FT TURN 404 405

Query Match 53.8%; Score 49; DB 1; Length 492;
Best Local Similarity 83.3%; Pred. No. 8.7;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 WFWFW 10

```
Db 6 WRWEPW 11
|||||
RESULT 4
ID_YD55 MYCTU STANDARD; PRT; 715 AA.
AC Q11025;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein Rv1355c/WT1398.
GN Rv1355C OR M1398 OR MTCY02B10.19C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RL laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z79700; CAB02005.1; -.
CC EMBL; AE006982; AAK45219.1; -.
CC PIR; G70715; G70715.
CC TIGR; MT0971; -.
CC Tuberculist; Rv0945; -.
CC InterPro; IPR002198; ADH_short.
CC Pfam; PF00106; adh_short; 1.
CC PRINTS; PR00080; SDRFAMILY.
CC PROSITE; PS00061; ADH_SHORT; 1.
KW Hypothetical protein; Complete proteome.
FT ACT SITE 159 159 BY SIMILARITY.
SQ SEQUENCE 715 AA; 78181 MW; 455495248A56041C CRC64;
Query Match 53.8%; Score 49; DB 1; Length 715;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 3 KXWPPWPPRR 12
|:|:|:|:|
Db 64 KWAYYPWRR 73
```

```
RESULT 5
ID_Y945 MYCTU STANDARD; PRT; 253 AA.
AC P71564;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative oxidoreductase Rv0945/MT0971 (EC 1.-.-.-).
GN Rv0945 OR MT0971 OR MTCY10D7.29C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RL laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z79700; CAB02005.1; -.
CC EMBL; AE006982; AAK45219.1; -.
CC PIR; G70715; G70715.
CC TIGR; MT0971; -.
CC Tuberculist; Rv0945; -.
CC InterPro; IPR002198; ADH_short.
CC Pfam; PF00106; adh_short; 1.
CC PRINTS; PR00080; SDRFAMILY.
CC PROSITE; PS00061; ADH_SHORT; 1.
KW Hypothetical protein; Oxidoreductase; Complete proteome.
FT ACT SITE 159 159 BY SIMILARITY.
SQ SEQUENCE 253 AA; 27138 MW; BAD937208842DA12 CRC64;
Query Match 51.8%; Score 47; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 8; 7;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 PWWPW 10
|:|:|:|:|
Db 230 PWWPW 234
```

```

RESULT 6
VGL2_IBVD2
ID VGL2_IBVD2 STANDARD; PRT; 1154 AA.
AC P12722; Q66176; Q66177;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein)
DE [Contains: Spike protein S1; Spike protein S2].
GN S.

OS Avian infectious bronchitis virus (strain D274) (IBV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11124;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89386000; PubMed=2550899;
RT Jordi B.J.A.M., Kremers D.A.W.M., Kusters H.G., van der Zeijst B.A.M.;
RT "Nucleotide sequence of the gene coding for the peplomer protein (=
RT spike protein) of infectious bronchitis virus, strain D274.";
RL Nucleic Acids Res. 17:6726-6726(1989)
CC -!- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS
CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X15832; CAA33837.1; -.
PIR; A34300; VGIHIB.
DR InterPro; IPR002551; Corona_S1.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01600; Corona_S1; 1.
DR Pfam; PF01601; Corona_S2; 1.
KW Glycoprotein; Envelope protein; Transmembrane; Signal.
FT CHAIN 1 18
FT CHAIN 19 1154 E2 GLYCOPROTEIN.
FT CHAIN 19 538 SPIKE PROTEIN S1.
FT CHAIN 539 1154 SPIKE PROTEIN S2.
FT CHAIN 1121 1138 CYS-RICH.
FT DOMAIN 23 23
FT CARBOHYD 23 23 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 74 74 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 139 139 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 164 164 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 213 213 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 248 248 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 514 514 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 543 543 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 592 592 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 677 677 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 948 948 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 961 961 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1015 1015 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1052 1052 N-LINKED (GLCNAC. .) (POTENTIAL).

RESULT 7
VGL2_IBVB
ID VGL2_IBVB STANDARD; PRT; 1162 AA.
AC P11223; P05134;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein)
DE [Contains: Spike protein S1; Spike protein S2].
GN S.

OS Avian infectious bronchitis virus (strain Beaudette) (IBV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11122;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85159540; PubMed=2984314;
RT Binns M.M., Boursnell M.E.G., Cavanagh D., Pappind D.J.C.,
RT Brown T.D.K.;
RT "Cloning and sequencing of the gene encoding the spike protein of the
RT coronavirus IBV.";
RL J. Gen. Virol. 66:719-726(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87085499; PubMed=3025348;
RT Binns M.M., Boursnell M.E.G., Tomley F.M., Brown T.D.K.;
RT "Comparison of the spike precursor sequences of coronavirus IBV
RT strains M41 and 6/82 with that of IBV Beaudette.";
RL J. Gen. Virol. 67:2825-2831(1986).
CC -!- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS
CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; M95169; AAA70235.1; -.
EMBL; X02342; CAA26201.1; -.
PIR; S14939; S14939.
DR InterPro; IPR002551; Corona_S1.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01600; Corona_S1; 1.
DR Pfam; PF01601; Corona_S2; 1.
KW Glycoprotein; Envelope protein; Transmembrane; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1162 E2 GLYCOPROTEIN.
FT CHAIN 19 537 SPIKE PROTEIN S1.
FT CHAIN 538 1162 SPIKE PROTEIN S2.
FT CHAIN 1120 1137 CYS-RICH.
FT DOMAIN 51 51
FT CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 178 178 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).

```

FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 254 264 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 513 513 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 676 676 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 947 947 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 960 960 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 979 979 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1014 1014 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1038 1038 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1162 AA; 128046 MW; 0BAAD58113C8BED5 CRC64;
 Query Match 51.1%; Score 46.5; DB 1; Length 1162;
 Best Local Similarity 61.5%; Pred. No. 42;
 Matches 8; Conservative 1; Mismatches 1; Indels 3; Gaps 1;
 Qy 1 ILK---KWPWPW 10
 Db 1085 ILKYIKWPWYVW 1097
 RESULT 8
 VGL2 IBVK STANDARD; PRT; 1162 AA.
 ID VGL2 IBVK
 AC P12650;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE E2 glycoprotein precursor (Spike glycoprotein) (Peplover protein)
 DE [Contains: Spike protein S1; Spike protein S2].
 GN S.
 OS Avian infectious bronchitis virus (strain KB8523) (IBV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=11126;
 RN [1]
 RP SEQUENCE FROM N.A.
 EX MEDLINE=88306251; PubMed=2841803;
 RA Sutou S., Sato S., Okabe T., Nakai M., Sasaki N.;
 RT "Cloning and sequencing of genes encoding structural proteins of
 RT avian infectious bronchitis virus.";
 RT Virology 165:589-595(1988).
 CC -!- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS
 CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M21515; AAA6578.1; -;
 DR PIR; B29249; VGIHAK.
 DR InterPro; IPR002551; Corona S1.
 DR InterPro; IPR002552; Corona S2.
 DR Pfam; PF01600; Corona S1; 1
 DR Pfam; PF01601; Corona S2; 1
 DR Pfam; PF01601; Corona S2; 1
 KW Glycoprotein; Envelope protein; Transmembrane; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 1162 E2 GLYCOPROTEIN.
 FT CHAIN 19 537 SPIKE PROTEIN S1.
 FT CHAIN

FT CHAIN 538 1162 SPIKE PROTEIN S2.
 FT DOMAIN 1120 1137 CYS-RICH.
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 264 264 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 513 513 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 676 676 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 947 947 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 960 960 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 979 979 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1014 1014 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1058 1058 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1162 AA; 128537 MW; 2299036B35978A9F CRC64;
 Query Match 51.1%; Score 46.5; DB 1; Length 1162;
 Best Local Similarity 61.5%; Pred. No. 42;
 Matches 8; Conservative 1; Mismatches 1; Indels 3; Gaps 1;
 Qy 1 ILK---KWPWPW 10
 Db 1085 ILKYIKWPWYVW 1097
 RESULT 9
 VGL2 IBVM STANDARD; PRT; 1162 AA.
 ID VGL2 IBVM
 AC P12651;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE E2 glycoprotein precursor (Spike glycoprotein) (Peplover protein)
 DE [Contains: Spike protein S1; Spike protein S2].
 GN S.
 OS Avian infectious bronchitis virus (strain M41) (IBV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=11127;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87021475; PubMed=2429473;
 RA Niesters H.G.M., Lenstra J.A., Spaan W.J.M., Zijderveld A.J.,
 RA Bleumink-Pluym N.M.C., Hong F., van Scharrenburg G.J.M.,
 RA Horzinek M.C., van der Zeijst B.A.M.;
 RT "The peplover protein sequence of the M41 strain of coronavirus IBV
 RT and its comparison with Beaudette strains.";
 RL Virus Res. 5:253-263(1986).
 CC -!- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS
 CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M21515; AAA6578.1; -;
 DR PIR; B29249; VGIHAK.
 DR InterPro; IPR002551; Corona S1.
 DR InterPro; IPR002552; Corona S2.
 DR Pfam; PF01600; Corona S1; 1
 DR Pfam; PF01601; Corona S2; 1
 DR Pfam; PF01601; Corona S2; 1
 KW Glycoprotein; Envelope protein; Transmembrane; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 1162 E2 GLYCOPROTEIN.
 FT CHAIN 19 537 SPIKE PROTEIN S1.
 FT CHAIN

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
CC EMBL; M21883; AAA66575.1; -.
CC EMBL; A24863; CAA01736.1; -.
CC PIR; S07421; S07421.
CC InterPro; IPR002551; Corona_S1.
CC InterPro; IPR002552; Corona_S2.
CC Pfam; PF01600; Corona_S1; 1.
CC Pfam; PF01601; Corona_S2; 1.
CC Pfam; PF01601; Corona_S2; 1.
CC Glycoprotein; Envelope protein; Transmembrane; Signal.
CC SIGNAL 1 18
CC CHAIN 19 1163 E2 GLYCOPROTEIN.
CC CHAIN 19 537 SPIKE PROTEIN S1.
CC CHAIN 539 1163 SPIKE PROTEIN S2.
CC DOMAIN 1120 1137 CYS-RICH.
CC FT CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 178 178 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 212 212 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 247 247 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 264 264 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 271 271 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 425 425 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 513 513 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 530 530 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 579 579 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 591 591 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 599 599 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 676 676 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 714 714 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 947 947 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 960 960 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 979 979 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 1014 1014 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 1038 1038 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SQ SEQUENCE 1162 AA; 128077 MW; 3C9CC70938492DDA CRC64;

Query Match 51.1%; Score 46.5; DB 1; Length 1162;
Best Local Similarity 61.5%; Pred. No. 42;
Matches 8; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 1 ILK---KWPWPWP 10
Db 1085 ILKTYIKWPWPVW 1097

RESULT 10
VGL2_IBV6 STANDARD; PRT; 1163 AA.
AC P05135;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein)
DE [Contains: Spike protein S1; Spike protein S2].
GN S.
OS Avian infectious bronchitis virus (strain 6/82) (IBV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11121;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87085499; PubMed=3025348;
```

```
RA Binns M.M., Boursnell M.E.G., Tomley F.M., Brown T.D.K.;
RT Comparison of the spike precursor sequences of coronavirus IBV
RL strains M41 and 6/82 with that of IBV Beaudette.;
RL J. Gen. Virol. 67:2825-2831(1986).
CC -I- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS
CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
CC EMBL; X04723; CAA28432.1; -.
CC InterPro; IPR002551; Corona_S1.
CC InterPro; IPR002552; Corona_S2.
CC Pfam; PF01600; Corona_S1; 1.
CC Pfam; PF01601; Corona_S2; 1.
CC Pfam; PF01601; Corona_S2; 1.
CC Glycoprotein; Transmembrane; Signal.
CC SIGNAL 1 18
CC CHAIN 19 1163 E2 GLYCOPROTEIN.
CC CHAIN 19 537 SPIKE PROTEIN S1.
CC CHAIN 539 1163 SPIKE PROTEIN S2.
CC DOMAIN 1121 1138 CYS-RICH.
CC FT CARBOHYD 23 23 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 74 74 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 139 139 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 164 164 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 179 179 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 213 213 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 248 248 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 265 265 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 426 426 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 514 514 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 543 543 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 580 580 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 592 592 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 670 670 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 677 677 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 948 948 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 961 961 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 980 980 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 1015 1015 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 1052 1052 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 1075 1075 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SQ SEQUENCE 1163 AA; 128694 MW; 8FE344CF2995478C CRC64;

Query Match 51.1%; Score 46.5; DB 1; Length 1163;
Best Local Similarity 61.5%; Pred. No. 42;
Matches 8; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 1 ILK---KWPWPWP 10
Db 1086 ILKTYIKWPWPVW 1098

RESULT 11
YA05_SCHPO STANDARD; PRT; 136 AA.
ID YA05_SCHPO
AC Q09677;
DT 01-NOV-1995 (Rel. 32, Created)
```


RN SEQUENCE FROM N.A.
 RP STRAIN=H3TRV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skellon S., Squares S., Squares R.,
 RA Sultun J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN
 RP SEQUENCE FROM N.A.
 RX STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=22206494; PubMed=12218036;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Ustebach T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bisnai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
 RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL J. Bacteriol. 184:5479-5490(2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the mmpL family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z74020; CAA98334.1; -;
 CC EMBL; AE007027; AAK45875.1; -;
 CC FIR; B70763; B70763.
 CC TIGR; MT1608; -;
 CC TubercuList; RV1557; -;
 CC InterPro: IPR004869; MPEL.
 CC Pfam; PF03176; MPEL; 1.
 CC Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 161 181 POTENTIAL.
 FT TRANSMEM 190 210 POTENTIAL.
 FT TRANSMEM 214 234 POTENTIAL.
 FT TRANSMEM 242 262 POTENTIAL.
 FT TRANSMEM 293 313 POTENTIAL.
 FT TRANSMEM 330 350 POTENTIAL.
 SQ SEQUENCE 397 AA; 42421 MW; 678DC86E24472BF4 CRC64;
 Query Match 49.5%; Score 45; DB 1; Length 397;
 Best Local Similarity 54.5%; Pred. No. 24;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ILKKWFWNWR 11
 : : : : :
 Db 348 LLGRWFWNWR 358
 RESULT 14
 TRPE_PSESS STANDARD; PRT; 505 AA.
 ID TRPE_PSESS
 AC P21689;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Anthranilate synthase component I (EC 4.1.3.27).
 GN TRPE.
 OS Pseudomonas syringae (pv. savastanoi).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OC NCBI_TaxID=29438;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9100331; PubMed=1987141;
 RA da Costa e Silva O., Kosuge T.;
 RA "Molecular characterization and expression analysis of the
 RT anthranilate synthase gene of Pseudomonas syringae subsp.
 RT savastanoi.";
 RL J. Bacteriol. 173:463-471(1991).
 CC -1- CATALYTIC ACTIVITY: Chorismate + L-glutamine = anthranilate +
 CC Pyruvate + L-glutamate.
 CC -1- PATHWAY: Tryptophan biosynthesis; first step.
 CC -1- SUBUNIT: Tetramer of two components I and two components II (By
 CC similarity).
 CC -1- MISCELLANEOUS: Component I catalyzes the formation of anthranilate
 CC using ammonia rather than glutamine, whereas component II provides
 CC glutamine amidotransferase activity.
 CC -1- SIMILARITY: Belongs to the anthranilate synthase component I
 CC family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M55911; AAA26016.1; -;
 CC HSP; Q06128; LODL.
 CC InterPro: IPR005801; Anth_synth_chor.
 CC InterPro: IPR006805; Anth_synth_I_N.
 CC InterPro: IPR005256; Anth_synth.
 CC Pfam; PF04715; Anth_synth_I_N; 1.
 CC Pfam; PF00425; chorismate bind; 1.
 CC PRINTS; PR00095; ANTSNTHASEI.
 CC ProDom; PD000779; Anth_synth_chor; 1.
 CC TIGRPFAMS; TIGR00564; tpe_mst; 1.
 CC Tryptophan biosynthesis; Lyase.
 KW Tryptophan biosynthesis; Lyase.
 SQ SEQUENCE 505 AA; 56084 MW; A38E8193131F6BB CRC64;
 Query Match 49.5%; Score 45; DB 1; Length 505;
 Best Local Similarity 71.4%; Pred. No. 31;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 7 WWPWRK 13
 : : : : :
 Db 485 WWPWR 491
 RESULT 15
 FEN2_YEAST STANDARD; PRT; 512 AA.
 ID FEN2_YEAST
 AC P25621;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Probable transporter FEN2.
 GN FEN2 OR YCR028C OR YCR28C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OC NCBI_TaxID=4932;
 RN
 RP SEQUENCE FROM N.A.
 RA Cederberg H., Hohmann S., Schaaff-Gerstenschlaeger I., Huse K.,
 RA Zimmermann F.K.;
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93070619; PubMed=1332309;

RA Carbone M.L.A., Panzeri L., Falconi M.M., Carcano C., Plevani P.,
 RA Lucchini G.;
 RT "Nucleotide sequence of 9.2 kb left of CRY1 on yeast chromosome III
 RT from strain AS972: evidence for a Ty insertion and functional
 RT analysis of open reading frame YCR28.";
 RL Yeast 8:805-812(1992).
 RN [3]
 RN SIMILARITY TO DAL5 FAMILY.
 RP MEDLINE=94147996; PubMed=8313894;
 RX Koonin E.V., Bork P., Sander C.;
 RA "Yeast chromosome III: new gene functions.";
 RT EMBO J. 13:493-503(1994).
 RL [4]
 RN CHARACTERIZATION.
 RP MEDLINE=96367594; PubMed=871708;
 RX Marcireau C., Joets J., Pousset D., Guilloton M., Karst F.;
 RA "FEN2: a gene implicated in the catabolite repression-mediated
 RT regulation of ergosterol biosynthesis in yeast.";
 RL Yeast 12:531-539(1996).
 CC -!- FUNCTION: Involved in the catabolite repression-mediated
 CC regulation of ergosterol biosynthesis and in fenpropimorph
 CC resistance.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the allantate permease family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X59720; CAA42320.1; -.
 DR PIR; S19439; S19439.
 DR GertOnline; 138935; -.
 DR SGD; S0000623; FEN2
 DR GO; GO:0005886; C:plasma membrane; IGI.
 DR GO; GO:0015233; F:ipantothenate transporter activity; IGI.
 DR GO; GO:0015887; P:ipantothenate transport; IGI.
 KW Transmembrane; Transport.
 FT TRANSMEM 28 48 POTENTIAL.
 FT TRANSMEM 80 100 POTENTIAL.
 FT TRANSMEM 103 123 POTENTIAL.
 FT TRANSMEM 133 153 POTENTIAL.
 FT TRANSMEM 165 185 POTENTIAL.
 FT TRANSMEM 199 219 POTENTIAL.
 FT TRANSMEM 272 292 POTENTIAL.
 FT TRANSMEM 313 333 POTENTIAL.
 FT TRANSMEM 343 363 POTENTIAL.
 FT TRANSMEM 373 393 POTENTIAL.
 FT TRANSMEM 402 422 POTENTIAL.
 FT TRANSMEM 435 455 POTENTIAL.
 FT CONFLICT 104 104 W -> V (IN REF. 2).
 SQ SEQUENCE 512 AA; 58256 MW; 361942E74C62B3B4 CRC64;

Query Match 49.5%; Score 45; DB 1; Length 512;
 Best Local Similarity 62.5%; Pred. No. 31;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ILKKWFW 8
 :||:|
 Db 268 VLKRWFW 275

Search completed: May 4, 2004, 15:20:16
 Job time : 9.21053 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:14:57 ; Search time 34.5526 Seconds
(without alignments)
118.710 Million cell updates/sec

Title: US-09-444-281-35

Perfect score: 91

Sequence: 1 ILKKWPWPWRK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mbc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	62.6	723	12 Q9DUC4	Q9duc4 tt virus. o
2	54	59.3	137	10 Q84ST7	Q84st7 oryza sativ
3	54	59.3	1383	12 Q84712	Q84712 porcine epi
4	54	59.3	1383	12 Q91AV1	Q91av1 porcine epi
5	54	59.3	1383	12 Q8B482	Q8b482 porcine epi
6	54	59.3	1386	12 Q8Q998	Q8q998 porcine epi
7	52	57.1	102	16 Q8P429	Q8p429 xanthomonas
8	52	57.1	105	16 Q8PPU5	Q8ppu5 xanthomonas
9	52	57.1	351	16 Q8DJH5	Q8djh5 synchococ
10	52	57.1	745	12 Q9JH31	Q9jh31 tt virus. o
11	52	57.1	1018	17 Q9HKX3	Q9hxx3 thermoplasm
12	51	56.0	298	17 Q8ZU59	Q8zu59 pyrobaculum
13	51	56.0	299	4 Q9Y4N1	Q9y4n1 homo sapien
14	50.5	55.5	225	10 Q84ZR3	Q84zr3 oryza sativ
15	50	54.9	327	10 Q9AUN3	Q9aun3 oryza sativ
16	50	54.9	327	10 Q7XFD1	Q7xfd1 oryza sativ

ALIGNMENTS

RESULT 1

Q9DUC4 PRELIMINARY; PRT; 723 AA.
ID Q9DUC4
AC Q9DUC4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE ORF1.
OS TT virus.
OC Viruses; ssDNA viruses; Circoviridae; Anellovirus.
OX NCBI_TaxID=68887;
RN [1]
RP STRAIN=Mf-TTV9;
RC SEQUENCE FROM N.A.
RA Okamoto H.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Mf-TTV9;
RX MEDLINE=20534983; PubMed=11080484;
RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,
RA Kishimoto J., Iinaka T., Miyakawa Y., Mayumi M.;
RT "Species-specific TT viruses in humans and nonhuman primates and their phylogenetic relatedness.";
RL Virology 277:368-378(2000).
DR GO; AB041959; BAB19313.1;
DR GO; GO:0004185; P:serine carboxypeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001563; Peptidase_S10.
DR Pfam; PF02956; TT ORF1; 1.
DR PROSITE; PS00131; CARBOXYPEPT SER SER; 1.
SQ SEQUENCE 723 AA; 85393 MW; -232D003098766344 CRC64;

Query Match 62.6%; Score 57; DB 12; Length 723;

Best Local Similarity 100.0%; Pred.No. 9.4;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PWKPPRR 12

Db 2 PWPWRR 8

```

RESULT 2
Q84ST7 PRELIMINARY; PRT; 137 AA.
AC Q84ST7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein OSJNB0092N01.27.
GN OSJNB0092N01.27.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Isitrin T., Kim M.M., Bera J.O., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feidblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RA "Oryza sativa chromosome 3 BAC OSJNB0092N01 genomic sequence.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC120535; AAO73229.1; -
KW Hypothetical protein.
SQ SEQUENCE 137 AA; 15795 MW; 67862CFAD153CE99 CRC64;

Query Match 59.3%; Score 54; DB 10; Length 137;
Best Local Similarity 50.0%; Pred. No. 5.2;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILKWPWRR 12
Db 51 LVRRPWPWRR 62

RESULT 3
Q84712 PRELIMINARY; PRT; 1383 AA.
AC Q84712;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Spike protein.
OS Porcine epidemic diarrhea virus (strain Br1/87) (PEDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=229033;
RN [1]
RN SEQUENCE FROM N.A.
RC MEDLINE=94231173; PubMed=8176382;
RA Duarte M., Laude H.;
RA "Sequence of the spike protein of the porcine epidemic diarrhoea
RT virus.";
RL J. Gen. Virol. 75:1195-1200(1994).
RN [2]
RN SEQUENCE FROM N.A.
RC MEDLINE=93389433; PubMed=8397280;
RA Bridgen A., Duarte M., Tobler K., Laude H., Ackermann M.;
RA "Sequence determination of the nucleocapsid protein gene of the
RT porcine epidemic diarrhoea virus confirms that this virus is a
coronavirus related to human coronavirus 229E and porcine
coronavirus related to human coronavirus 229E and porcine

```

```

transmissible gastroenteritis virus.";
RL J. Gen. Virol. 74:1795-1804(1993).
RN [3]
RN SEQUENCE FROM N.A.
RC MEDLINE=94120721; PubMed=8291230;
RA Duarte M., Tobler K., Bridgen A., Rasschaert D., Ackermann M.,
RA Laude H.;
RA "Sequence analysis of the porcine epidemic diarrhea virus genome
RT between the nucleocapsid and spike protein genes reveals a polymo.";
RL Virology 198:466-476(1994).
DR EMBL; Z25483; CAA80971.1; -.
DR InterPro; IPR002551; Corona_S1.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01600; Corona_S1; 1.
DR Pfam; PF01601; Corona_S2; 1.
FT CONFLICT 422 422
SQ SEQUENCE 1383 AA; 151405 MW; 741C84D5DD3BDC4D CRC64;

Query Match 59.3%; Score 54; DB 12; Length 1383;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 KWPWPPW 10
Db 1322 KWPWPPW 1328

RESULT 4
Q91AV1 PRELIMINARY; PRT; 1383 AA.
AC Q91AV1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Spike protein.
OS Porcine epidemic diarrhea virus (strain CV777) (PEDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=229032;
RN [1]
RN SEQUENCE FROM N.A.
RC MEDLINE=93389433; PubMed=8397280;
RA Bridgen A., Duarte M., Tobler K., Laude H., Ackermann M.;
RA "Sequence determination of the nucleocapsid protein gene of the
RT porcine epidemic diarrhoea virus confirms that this virus is a
coronavirus related to human coronavirus 229E and porcine
transmissible gastroenteritis virus.";
RL J. Gen. Virol. 74:1795-1804(1993).
RN [2]
RN SEQUENCE FROM N.A.
RC MEDLINE=94120721; PubMed=8291230;
RA Duarte M., Tobler K., Bridgen A., Rasschaert D., Ackermann M.,
RA Laude H.;
RA "Sequence analysis of the porcine epidemic diarrhea virus genome
RT between the nucleocapsid and spike protein genes reveals a polymorphic
ORF.";
RL Virology 198:466-476(1994).
RN [3]
RN SEQUENCE FROM N.A.
RC MEDLINE=96112302; PubMed=8830538;
RA Tobler K., Ackermann M.;
RA "PEDV leader sequence and junction sites.";
RL Adv. Exp. Med. Biol. 380:541-542(1995).
RN [4]
RN SEQUENCE FROM N.A.
RC MEDLINE=98455678; PubMed=9782358;
RA Bridgen A., Kocherhans R., Tobler K., Carvajal A., Ackermann M.;
RA "Further analysis of the genome of porcine epidemic diarrhoea virus.";
RL Adv. Exp. Med. Biol. 440:781-786(1998).
RN [5]
RN SEQUENCE FROM N.A.
RC MEDLINE=98455678; PubMed=9782358;
RA Kocherhans R., Bridgen A., Ackermann M., Tobler K.;
RA "The complete genome sequence of porcine epidemic diarrhoea
RT

```

```
RT   coronavirus ";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DBSJ databases.
DR   EMBL; AF353511; AAK38656.1; -.
DR   InterPro; IPR002551; Corona_S1.
DR   InterPro; IPR002552; Corona_S2.
DR   Pfam; PF01600; Corona_S1; 1.
DR   Pfam; PF01601; Corona_S2; 1.
SQ   SEQUENCE 1383 AA; 151352 MW; 02285E5E5435876D CRC64;

Query Match      59.3%; Score 54; DB 12; Length 1383;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 KWPWWPW 10
Db      1322 KWPWWPW 1328

RESULT 5
QB482
ID      QB482      PRELIMINARY;      PRT; 1383 AA.
AC      QB482;
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Spike protein.
OS      Porcine epidemic diarrhea virus.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC      Coronaviridae; Coronavirus.
OX      NCBI_TaxID=28295;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Chinju99;
RA      Yeo S.-G., Krell P., Nagy E.;
RT      "Cloning and nucleotide sequence analysis of spike gene of porcine
RL      epidemic diarrhea virus detected in Korea.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DBSJ databases.
DR   EMBL; AY167585; AAN86621.1; -.
DR   InterPro; IPR002551; Corona_S1.
DR   InterPro; IPR002552; Corona_S2.
DR   Pfam; PF01600; Corona_S1; 1.
DR   Pfam; PF01601; Corona_S2; 1.
SQ   SEQUENCE 1383 AA; 151582 MW; B5BA4D7EE5371A54 CRC64;

Query Match      59.3%; Score 54; DB 12; Length 1383;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 KWPWWPW 10
Db      1322 KWPWWPW 1328

RESULT 6
QB482
ID      QB482      PRELIMINARY;      PRT; 1386 AA.
AC      QB482;
DT      01-JUN-2002 (TrEMBLrel. 21, Created)
DT      01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT      01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE      Spike protein.
OS      Porcine epidemic diarrhea virus.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC      Coronaviridae; Coronavirus.
OX      NCBI_TaxID=28295;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Kang T.-J., Lim Y.-Y., Jang Y.-S., Kwon T.-H., Kim D.-H., Yang M.-S.;
RT      "Spike protein gene of Korea Porcine Epidemic Diarrhea Virus.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DBSJ databases.
DR   EMBL; AF500215; AAM19716.1; -.
DR   InterPro; IPR002551; Corona_S1.

RT   coronavirus ";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DBSJ databases.
DR   EMBL; AF353511; AAK38656.1; -.
DR   InterPro; IPR002551; Corona_S1.
DR   InterPro; IPR002552; Corona_S2.
DR   Pfam; PF01600; Corona_S1; 1.
DR   Pfam; PF01601; Corona_S2; 1.
SQ   SEQUENCE 1386 AA; 151853 MW; 11F98ECB2AA0526B CRC64;

Query Match      59.3%; Score 54; DB 12; Length 1386;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 KWPWWPW 10
Db      1325 KWPWWPW 1331

RESULT 7
QB429
ID      QB429      PRELIMINARY;      PRT; 102 AA.
AC      QB429;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE      Inner membrane protein.
OS      Xanthomonas campestris (pv. campestris).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC      Xanthomonadaceae; Xanthomonas.
OX      NCBI_TaxID=340;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 33913 / NCPPB 528;
RC      MEDLINE=22022145; PubMed=12024217;
RA      da Silva A.C.B., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA      Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA      Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA      Camarotte G., Canavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA      Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA      Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA      Formighieri E.F., Franco M.C., Greggio C.C., Lemos E.G.M., Lemos M.V.F.,
RA      Katsuyama A.M., Kishi L.T., Leite R.P., Madsen A.M.B.N., Martinez-Rossi N.M.,
RA      Locati E.C., Machado M.A., Madeira A.M.B.N., Miyaki C.Y., Moon D.H.,
RA      Martins E.C., Meidanis J., Menck C.F.M., Oliveira M.C., Oliveira V.R.,
RA      Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA      Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA      Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA      Trindade dos Santos M., Truffi D., Tsei S.M., White F.F.,
RA      Setubal J.C., Kitajima J.P.;
RT      "Comparison of the genomes of two Xanthomonas pathogens with differing
RT      host specificities.";
RL   Nature 417:459-463(2002).
DR   EMBL; AE012475; AAM42819.1; -.
KW      Complete proteome.
SQ   SEQUENCE 102 AA; 11488 MW; 641654465C9571BF CRC64;

Query Match      57.1%; Score 52; DB 16; Length 102;
Best Local Similarity 71.4%; Pred. No. 7.3;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      4 KWPWWPW 10
Db      65 RWPWWAW 71

RESULT 8
QBPPUS
ID      QBPPUS      PRELIMINARY;      PRT; 105 AA.
AC      QBPPUS;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE      Inner membrane protein.
OS      Xanthomonas axonopodis (pv. citri).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
```

```
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=1204217;
da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chamberg F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.P., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.F.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
EMBL: AB011686; AM35479.1; -.
KW Complete proteome.
SQ SEQUENCE 105 AA; 11853 MW; 4DF5A59FBC5EF3C2 CRC64;

Query Match 57.1%; Score 52; DB 16; Length 105;
Best Local Similarity 71.4%; Pred. No. 7.4;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 KWPWPWP 10
Db 65 RWPWPWP 71

RESULT 9
Q8DJH5 PRELIMINARY; PRT; 351 AA.
AC Q8DJH5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Tlr1250 protein.
DE Tlr1250.
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iiguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL: AP005373; BAC08802.1; -.
DR InterPro; IPR001679; Sun_Nop1/Nop2.
DR Pfam; PF01189; Nop1_Nop2_Sun; 1.
KW Complete proteome.
SQ SEQUENCE 351 AA; 38494 MW; 675046ADCBE7C935 CRC64;

Query Match 57.1%; Score 52; DB 16; Length 351;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KWPWPWP 9
Db 2 KWPWPWP 7

RESULT 10
Q9JH31 PRELIMINARY; PRT; 746 AA.
AC Q9JH31;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE ORF1.
OS TT virus.
OC Viruses; ssDNA viruses; Circoviridae; Anellovirus.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TJN02;
RA Okamoto H.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TJN02;
RX MEDLINE=20436801; PubMed=11003468;
RA Ukita M., Okamoto H., Nishizawa T., Tawara A., Takahashi M.,
RA Iizuka H., Miyakawa Y., Mayumi M.;
RT "The entire nucleotide sequences of two distinct TT virus (TTV)
RT isolates (TJN01 and TJN02) remotely related to the original TTV
RT isolates.";
RL Arch. Virol. 145:1543-1559(2000).
DR EMBL: AB028669; BAA94878.1; -.
DR InterPro; IPR004219; TTVirus_Unk.
DR Pfam; PF02956; TT_ORF1; 1.
SQ SEQUENCE 746 AA; 88561 MW; B0B22953AE764E3E CRC64;

Query Match 57.1%; Score 52; DB 12; Length 746;
Best Local Similarity 66.7%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 WPMWPERK 13
Db 3 WGMWRWERR 11

RESULT 11
Q9HKX3 PRELIMINARY; PRT; 1018 AA.
AC Q9HKX3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Conserved hypothetical membrane protein.
DE TA0470.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum";
RL Nature 407:508-513(2000).
DR EMBL: AL445054; CAC11612.1; -.
DR InterPro; IPR000731; SSD_5TM.
DR PROSITE; PS0156; SSD; 1.
KW Complete proteome.
SQ SEQUENCE 1018 AA; 112322 MW; 83BE84D33C74B852 CRC64;

Query Match 57.1%; Score 52; DB 17; Length 1018;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 ILKWPWWP 9
Db 1004 LMKWNP 1012

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 12
Q8ZU59 PRELIMINARY; PRT; 298 AA.
AC Q8ZU59;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dihydropterolate synthase.
GN PAE2937.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51769 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009902; AAL64549.1; -.
DR GO; GO:0004156; P:dihydropterolate synthase activity; IEA.
DR GO; GO:0009396; P:folic acid and derivative biosynthesis; IEA.
DR InterPro; IPR000489; Dhdropt_synth.
DR InterPro; IPR006390; DHPS.
DR Pfam; PF00809; Pterin bind; 1.
DR TIGRFAMs; TIGR01496; DHPS; 1.
KW Complete proteome.
SQ SEQUENCE 298 AA; 32885 MW; 0A463F36739D3ED1 CRC64;

Query Match 56.0%; Score 51; DB 17; Length 298;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 KWPWWP 10
Db 209 QWFWKW 215

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

RESULT 13
Q9Y4N1 PRELIMINARY; PRT; 299 AA.
AC Q9Y4N1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZF434C192.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue: testis;
RA Ansong W., Winkler U., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL096753; CAB46428.2; -.
DR PIR; T12505; T12505.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 299 AA; 34032 MW; 6B8DB60E6A88239A CRC64;

Query Match 56.0%; Score 51; DB 4; Length 299;
Best Local Similarity 85.7%; Pred. No. 26;

QY 6 PWPWRR 12
Db 37 PWPWRR 43

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 14
Q84ZK3 PRELIMINARY; PRT; 225 AA.
AC Q84ZK3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE OJ1372_D12.7 protein.
GN OJ1372_D12.7.
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RT clone: OJ1372_D12.7";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003827; BAC57651.1; -.
SQ SEQUENCE 225 AA; 23825 MW; 52096C5EA0083F77 CRC64;

Query Match 55.5%; Score 50.5; DB 10; Length 225;
Best Local Similarity 53.8%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY 4 KWPW---WPPRRK 13
Db 111 RWCWAAFPWRRR 123

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 15
Q9AUN3 PRELIMINARY; PRT; 327 AA.
AC Q9AUN3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSUNBA0058E19.18.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Spiegel L.A., King L., Kirchoff K.A., de la Bastide M., Preston R.R.,
RA Nacimento L.U., Vil M.D., Baker J.P., Miller B., Cummins D.M.,
RA Kuit K.H., Rodriguez S., Santos L., Zutavern T., Ballia V.S.,
RA Shah R.S., Bahret A., Bal H.P., O'Shaughnessy A., Dedhia N.N.,
RA McCombie W.R.;
RT "Genomic Sequence For Oryza sativa, Nipponbare Strain, Chromosome X,
RT Clone OSUNBA0058E19, Complete Sequence.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC089945; AAK13143.1; -.
DR Gramene; Q9AUN3; -.
KW Hypothetical protein.
SQ SEQUENCE 327 AA; 36672 MW; 5CCA9080664BDOCA CRC64;

Query Match 54.9%; Score 50; DB 10; Length 327;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 7 WWPERR 12
|||
Db 119 WWPERR 124

Search completed: May 4, 2004, 15:22:10
Job time : 35.886 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:08:11 ; Search time 45.7895 Seconds
(without alignments)
74.047 Million cell updates/sec

Title: US-09-444-281-36

Perfect score: 86

Sequence: 1 ILRPWPWPRRK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	100.0	12	2	AAy24550 Indolicid
2	86	100.0	12	3	AAy94496 MBI-11B7
3	86	100.0	12	3	AAy91791 Amino aci
4	86	100.0	12	6	ADA00524 Antimicro
5	86	100.0	12	7	ADC98871 Synthetic
6	86	100.0	12	7	ADC98903 Synthetic
7	86	100.0	14	7	ADC98990 Synthetic
8	86	100.0	14	7	ADC98989 Synthetic
9	86	100.0	20	2	AAy24553 Indolicid
10	86	100.0	20	3	AAy91797 Amino aci
11	86	100.0	20	6	ADA00530 Antimicro
12	86	100.0	20	7	ADC98878 Synthetic
13	86	100.0	21	2	AAW63376 Cationic
14	86	100.0	21	2	AAy24554 Indolicid
15	86	100.0	21	2	AAy24552 Indolicid
16	86	100.0	21	3	AAy91798 Amino aci
17	86	100.0	21	3	AAy91796 Amino aci
18	86	100.0	21	6	ADA00529 Antimicro
19	86	100.0	21	6	ADA00531 Antimicro
20	86	100.0	21	7	ADC98877 Synthetic
21	86	100.0	21	7	ADC98879 Synthetic
22	86	100.0	27	2	AAW63363 Indolicid
23	86	100.0	28	3	AAy91800 Amino aci
24	86	100.0	28	6	ADA00533 Antimicro
25	86	100.0	28	7	ADC98881 Synthetic

26	83	96.5	12	2	AAy24567 Indolicid
27	83	96.5	12	3	AAy91788 Amino aci
28	83	96.5	12	6	ADA00521 Antimicro
29	83	96.5	12	7	ADC98868 Synthetic
30	82	95.3	12	2	AAW63364 Indolicid
31	82	95.3	12	2	AAy24594 Indolicid
32	82	95.3	12	3	AAy91817 Amino aci
33	82	95.3	12	3	AAy91841 Amino aci
34	82	95.3	12	6	ADA00552 Antimicro
35	82	95.3	12	6	ADA00586 Antimicro
36	82	95.3	12	7	ADC98943 Synthetic
37	82	95.3	12	7	ADC98902 Synthetic
38	81	94.2	12	2	AAy24595 Indolicid
39	81	94.2	12	2	AAy24605 Indolicid
40	81	94.2	12	3	AAy91852 Amino aci
41	81	94.2	12	3	AAy91842 Amino aci
42	81	94.2	12	6	ADA00587 Antimicro
43	81	94.2	12	6	ADA00597 Antimicro
44	81	94.2	12	7	ADC98954 Synthetic
45	81	94.2	12	7	ADC98944 Synthetic

ALIGNMENTS

RESULT 1
AAy24550
ID AAY24550 standard; peptide; 12 AA.
XX
AC AAY24550;
XX
DT 18-AUG-1999 (first entry)
XX
DE Indolicidin analogue #2.
XX
KW Indolicidin; bacterial infection; photo-oxidised solubiliser;
KW antimicrobial; antibiotic; antihistaminic; surface disinfectant; additive;
KW shampoo; soap; insecticide; herbicide; preservative; food;
KW technical material.
XX
OS Synthetic.
XX
PN WO9807745-A2.
XX
PD 26-FEB-1998.
XX
PF 21-AUG-1997; 97WO-US014779.
XX
PR 21-AUG-1996; 96US-0024754P.
PR 13-JAN-1997; 97US-0034949P.
XX
PA (MICR-) MICROLOGIX BIOTECH INC.
XX
PI Fraser JR, West MH, Krieger TJ, Taylor R, Erfle D;
XX WPI; 1998-169090/15.
XX
DR New indolicidin analogues with antimicrobial activity and related nucleic
XX acid - vectors, transformed cells and antibodies, also conjugates with
XX polyoxalkylene glycol and fatty acid to reduce toxicity, useful
XX therapeutically, as disinfectants etc.
XX
PS Claim 11; Page 88; 129pp; English.
XX
CC AAY24549 to AAY24615 represent indolicidin analogues of formulae (I)-
CC (VIII) containing up to 25 amino acids (aa): RXXXXXB (I), BXXXXXB (II),
CC BXXXXXB (III), BXXXXXB (IV), BXXXXXB (V), BXXXXXB (VI), BXXXXXB (VII),
CC LBBXXXXXB (VIII), LBBXXXXXB (IX), LBBXXXXXB (X), LBBXXXXXB (XI),
CC P or V; X = hydrophobic residue, preferably W; B = basic aa, preferably R
CC or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V; in (VII) at
CC least 2 X = F or Y. The analogues are used to treat infections caused by
CC bacteria (Gram positive or negative, or anaerobic); fungi (yeast or
CC moulds); parasites (protozoa, nematodes, cestodes or trematodes) or

CC viruses. Typical of very many pathogens that can be controlled are
 CC Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola hepatica,
 CC Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus aureus,
 CC Listeria, Clostridium, rotavirus and papilloma virus. Compounds derived
 CC from the analogues may be used similarly; the compounds may also be
 CC prepared from antibiotics or antirhythmic agents. The analogues may be
 CC used therapeutically or to coat medical devices; also they are useful as
 CC surface disinfectants, as additives to shampoo or soaps, as insecticides
 CC or herbicides, or as preservatives for foods and technical materials. The
 CC analogues are administered by injection, lavage, orally or topically,
 CC generally at 0.1-50 mg/kg. These analogues have a broader spectrum of
 CC activity than indolicidin and modification as compounds reduces their
 CC toxicity
 CC XX
 CC SQ

Sequence 12 AA;

Query Match 100.0%; Score 86; DB 2; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILRWPWPWRRK 12
 |||||
 Db 1 ILRWPWPWRRK 12

RESULT 2
 AAY94496
 ID AAY94496 standard; peptide; 12 AA.

AC AAY94496;

DT 20-SEP-2000 (first entry)

DE MBI-11B7 peptide derived from indolicidin.

CC Cellulose binding domain; CBD; cationic peptide; MBI-11B7; indolicidin;
 CC bovine.

OS Bos taurus.

PN WO200031279-A2.

PD 02-JUN-2000.

PF 19-NOV-1999; 99WO-CA001107.

PR 20-NOV-1998; 98US-0109218P.

XX (MICR-) MICROLOGIX BIOTECH INC.

PI Burian J, Bartfeld D;

DR WPI; 2000-400086/34.

CC Multi-domain fusion protein expression cassette used for high yield
 CC stable production of foreign peptide gene products.

PS Disclosure; Page 24; 73pp; English.

CC A novel method allows the efficient production of cationic peptides in
 CC recombinant host cells. The method involves construction of a multi-
 CC domain fusion protein expression cassette comprising a promoter and a
 CC nucleic acid molecule expressed as an insoluble protein. The inclusion of
 CC anionic peptide sequences in the linker sequences neutralises the
 CC positive charge of the cationic peptide so that the charge of the fusion
 CC protein is controlled. This cassette allows high yield, stable production
 CC of the cationic peptide. Cationic peptides such as bovine indolicidin may
 CC be used as antimicrobial agents. The present sequence is the MBI-11B7
 CC peptide. MBI-11B7 is a cationic peptide derived from modifications of
 CC indolicidin
 CC XX
 CC SQ

Sequence 12 AA;

Query Match 100.0%; Score 86; DB 3; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILRWPWPWRRK 12
 |||||
 Db 1 ILRWPWPWRRK 12

RESULT 3

AAY91791

ID AAY91791 standard; peptide; 12 AA.

AC AAY91791;

DT 06-JUN-2000 (first entry)

DE Amino acid sequence of cationic peptide MBI 11B7CN.

CC Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 CC leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
 CC breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 CC multidrug resistance.

OS Synthetic.

PN WO9965506-A2.

PD 23-DEC-1999.

PF 14-JUN-1999; 99WO-CA000552.

PR 12-JUN-1998; 98US-00096541.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Friedland HD, Krieger TU, Taylor R, Erfle D, Fraser JR, West MRP;

XX WPI; 2000-223549/19.

CC Novel pharmaceutical composition containing optionally activated
 CC polyoxyalkylene-modified cationic peptides, useful for treating tumors.

PS Claim 1; Page 14; 94pp; English.

CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxyalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon

XX Sequence 12 AA;

Query Match 100.0%; Score 86; DB 3; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILRWPWPWRRK 12
 |||||
 Db 1 ILRWPWPWRRK 12

RESULT 4

ADA00524

ID ADA00524 standard; peptide; 12 AA.

AC ADA00524;

DT 06-NOV-2003 (first entry)

XX DE Antimicrobial cationic peptide 11B7CN.

XX DE antimicrobial; cationic; viscosity-increasing agent; solvent; buffer;

KW antibacterial; virucide; antiinflammatory; fungicide; protozoacide;

KW parasiticide; vulnery; dermatological; herbicide; insecticide;

KW infection; systemic infection; sepsis; acne; disinfectant; herbicide;

KW insecticide; silicone sealant.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 12 /label= amidated

FT WO2003015809-A2.

XX PD 27-FEB-2003.

XX PF 21-AUG-2002; 2002WO-US026525.

XX PR 21-AUG-2001; 2001US-0314232P.

PR 20-AUG-2002; 2002US-00225087.

XX PA (MICR-) MICROLOGIX BIOTECH INC.

XX PI Krieger TJ, Mcnicol PJ, Fraser JR;

XX DR WPI; 2003-332767/31.

XX PT Composition containing stabilized antimicrobial cationic protein, useful

PT for treating infections, particularly where associated with in-dwelling

PT devices.

XX PS Claim 47; Page 47; 30pp; English.

XX CC The present invention describes a composition (A) comprising an

CC antimicrobial cationic peptide (I), a viscosity-increasing agent (II) and

CC a solvent (III). Also described is a composition comprising (I), buffer

CC (IV) and (III). (I) has antibacterial, virucide, antiinflammatory,

CC fungicide, protozoacide, parasiticide, vulnery, dermatological,

CC herbicide and insecticide activities. (A) can be used to reduce the

CC population of microflora (eukaryotes, prokaryotes or viruses) at a target

CC site, particularly for treatment or prevention of infections. They can be

CC used to treat a wide range of systemic infections (e.g. sepsis) and for

CC topical treatment of wounds, but most especially can be used: (i) at

CC sites where medical devices have been, or will be, inserted into the body

CC (alternatively, they are used to treat the devices); and (ii) at sites on

CC the skin (particularly for treating acne) or the mucosa. The devices

CC treated are especially central venous, vascular dialysis, pulmonary

CC artery, peritoneal dialysis or umbilical catheters. They may also be used

CC as surface disinfectants; for treatment of clothing and air filters; in

CC cosmetics and soaps; as herbicides and insecticides; in building

CC materials (e.g. silicone sealants) and in processing animal products,

CC e.g. hides. The present sequence represents an antimicrobial cationic

CC peptide, which is used in the exemplification of the present invention.

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 86; DB 6; Length 12;

Best Local Similarity 100.0%; Pred. No. 1.1e-05; Mismatches 0; Gaps 0;

Matches 12; Conservative 0; Indels 0; Gaps 0;

QY 1 ILRWPPWPWRK 12

Db 1 ILRWPPWPWRK 12

RESULT 5

ADC98871

ID ADC98871 standard; peptide; 12 AA.

XX AC ADC98871;

XX DT 01-JAN-2004 (first entry)

XX DE Synthetic indolicidin analogue peptide - 11B7CN.

XX DE indolicidin analogue; antiseborrheic; dermatological; antiinflammatory;

KW antiarthritic; immunosuppressive; vulnery; antipruritic; antimicrobial;

KW antipruritic; neuroprotective; antipsoriatic; inflammation; acne;

KW arthritis; autoimmune disease; burn; Crohn's; colitis;

KW contact hypersensitivity; delayed; eczema; endotoxin shock syndrome;

KW fibromyositis; graft rejection; microbial infection; multiple sclerosis;

KW parapsoriasis; psoriasis; sclerosis; seborrhea.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "Optional N-terminal acetyl"

FT Modified-site 12 /note= "Optional C-terminal amide"

XX PN WO2003018619-A2.

XX PD 06-MAR-2003.

XX PF 26-AUG-2002; 2002WO-CA001351.

XX PR 24-AUG-2001; 2001US-0315003P.

PR 26-AUG-2002; 2002US-00229368.

XX PA (MICR-) MICROLOGIX BIOTECH INC.

PA (MCNI/) MCNICOL P J.

PA (PAWL/) PAWLAK S K.

PA (RUBI/) RUBINCHIK E.

PA (CAME/) CAMERON D.

PA (GUAR/) GUARNA M M.

XX PI Mcnicol PJ, Pawlak SK, Rubinchik E, Cameron D, Guarna MM;

XX DR WPI; 2003-393247/37.

XX PT Novel indolicidin analog useful for treating or preventing inflammation

PT at a target site associated with a condition such as acne, arthritis,

PT burn, Crohn's disease, colitis, and in image analysis and diagnostic

PT assays.

XX PS Example 1; Page 47; 66pp; English.

XX CC The invention relates to a novel indolicidin analogue. The analogue of

CC the invention demonstrates antiseborrheic, dermatological,

CC antiinflammatory, antiarthritic, immunosuppressive, vulnery,

CC antipruritic, antimicrobial, antipruritic, neuroprotective and

CC antipsoriatic and may be useful for treating or preventing inflammation

CC at a target site. The inflammation at the target site may be associated

CC with a condition selected from acne, arthritis, autoimmune disease, burn,

CC Crohn's disease, colitis, contact hypersensitivity, delayed

CC hypersensitivity, eczema, endotoxin shock syndrome, fibromyositis, graft

CC rejection, microbial infection, multiple sclerosis, parapsoriasis, of the

CC psoriasis, sclerosis and seborrhea. The current sequence is that of the

CC synthetic indolicidin analogue peptide of the invention.

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 86; DB 7; Length 12;

Best Local Similarity 100.0%; Pred. No. 1.1e-05; Mismatches 0; Gaps 0;

Matches 12; Conservative 0; Indels 0; Gaps 0;

QY 1 ILRWPPWPWRK 12

Db 1 ILRWPPWPWRK 12

RESULT 6

```

ADC98990
ID ADC98903 standard; peptide; 12 AA.
XX AC
XX AC ADC98903;
XX DT
XX 01-JAN-2004 (first entry)
XX DE
XX Synthetic indolicidin analogue peptide - 11D21CN.
XX indolicidin analogue; antiseborrheic; dermatological; antiinflammatory;
XX antiarthritic; immunosuppressive; vulnery; antipruritic; antimicrobial;
XX antipruritic; neuroprotective; antipsoriatic; inflammation; acne;
XX arthritis; autoimmune disease; burn; Crohn's; colitis;
XX contact hypersensitivity; delayed; eczema; endotoxin shock syndrome;
XX fibromyositis; graft rejection; microbial infection; multiple sclerosis;
XX parapsoriasis; psoriasis; sclerosis; seborrhea.
XX OS
XX Synthetic.
XX FH
XX Key Location/Qualifiers
XX Modified-site 14 /note= "C-terminal amide"
XX FT
XX WO2003018619-A2.
XX PN
XX 06-MAR-2003.
XX PD
XX 26-AUG-2002; 2002WO-CA001351.
XX PF
XX 24-AUG-2001; 2001US-0315003P.
XX PR
XX 26-AUG-2002; 2002US-00229368.
XX PA
XX (MICR-) MICROLOGIX BIOTECH INC.
XX (MCNI/) MCNICOL P J.
XX (PAWL/) PAWLAK S K.
XX (RUBI/) RUBINCHIK E.
XX (CAME/) CAMERON D.
XX (GUAR/) GUARNA M M.
XX PI
XX Mcnicol PJ, Pawlak SK, Rubinchik E, Cameron D, Guarna MM;
XX WPI; 2003-393247/37.
XX DR
XX Novel indolicidin analog useful for treating or preventing inflammation
XX at a target site associated with a condition such as acne, arthritis,
XX burn, Crohn's disease, colitis, and in image analysis and diagnostic
XX assays.
XX PS
XX Example 1; Page 48; 66pp; English.
XX CC
XX The invention relates to a novel indolicidin analogue. The analogue of
XX the invention demonstrates antiseborrheic, dermatological,
XX antiinflammatory, antiarthritic, immunosuppressive, vulnery,
XX antipruritic, antimicrobial, antipruritic, neuroprotective and
XX antipsoriatic and may be useful for treating or preventing inflammation
XX at a target site. The inflammation at the target site may be associated
XX with a condition selected from acne, arthritis, autoimmune disease, burn,
XX Crohn's disease, colitis, contact hypersensitivity, delayed
XX hypersensitivity, eczema, endotoxin shock syndrome, fibromyositis, graft
XX rejection, microbial infection, multiple sclerosis, parapsoriasis,
XX psoriasis, sclerosis and seborrhea. The current sequence is that of the
XX synthetic indolicidin analogue peptide of the invention.
XX SQ
XX Sequence 12 AA;
XX Query Match 100.0%; Score 86; DB 7; Length 12;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-05;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 ILRWPPWPPWRK 12
XX | | | | |
XX Db 1 ILRWPPWPPWRK 12
XX | | | | |
XX RESULT 7
XX ADC98990
XX ID ADC98990 standard; peptide; 14 AA.
XX AC
XX AC ADC98990;
XX DT
XX 01-JAN-2004 (first entry)
XX DE
XX Synthetic indolicidin analogue peptide - R11B7 H CN.
XX indolicidin analogue; antiseborrheic; dermatological; antiinflammatory;
XX antiarthritic; immunosuppressive; vulnery; antipruritic; antimicrobial;
XX antipruritic; neuroprotective; antipsoriatic; inflammation; acne;
XX arthritis; autoimmune disease; burn; Crohn's; colitis;
XX contact hypersensitivity; delayed; eczema; endotoxin shock syndrome;
XX fibromyositis; graft rejection; microbial infection; multiple sclerosis;
XX parapsoriasis; psoriasis; sclerosis; seborrhea.
XX OS
XX Synthetic.
XX FH
XX Key Location/Qualifiers
XX Modified-site 14 /note= "C-terminal amide"
XX FT
XX WO2003018619-A2.
XX PN
XX 06-MAR-2003.
XX PD
XX 26-AUG-2002; 2002WO-CA001351.
XX PF
XX 24-AUG-2001; 2001US-0315003P.
XX PR
XX 26-AUG-2002; 2002US-00229368.
XX PA
XX (MICR-) MICROLOGIX BIOTECH INC.
XX (MCNI/) MCNICOL P J.
XX (PAWL/) PAWLAK S K.
XX (RUBI/) RUBINCHIK E.
XX (CAME/) CAMERON D.
XX (GUAR/) GUARNA M M.
XX PI
XX Mcnicol PJ, Pawlak SK, Rubinchik E, Cameron D, Guarna MM;
XX WPI; 2003-393247/37.
XX DR
XX Novel indolicidin analog useful for treating or preventing inflammation
XX at a target site associated with a condition such as acne, arthritis,
XX burn, Crohn's disease, colitis, and in image analysis and diagnostic
XX assays.
XX PS
XX Example 1; Page 56; 66pp; English.
XX CC
XX The invention relates to a novel indolicidin analogue. The analogue of
XX the invention demonstrates antiseborrheic, dermatological,
XX antiinflammatory, antiarthritic, immunosuppressive, vulnery,
XX antipruritic, antimicrobial, antipruritic, neuroprotective and
XX antipsoriatic and may be useful for treating or preventing inflammation
XX at a target site. The inflammation at the target site may be associated
XX with a condition selected from acne, arthritis, autoimmune disease, burn,
XX Crohn's disease, colitis, contact hypersensitivity, delayed
XX hypersensitivity, eczema, endotoxin shock syndrome, fibromyositis, graft
XX rejection, microbial infection, multiple sclerosis, parapsoriasis,
XX psoriasis, sclerosis and seborrhea. The current sequence is that of the
XX synthetic indolicidin analogue peptide of the invention.
XX SQ
XX Sequence 14 AA;
XX Query Match 100.0%; Score 86; DB 7; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-05;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 ILRWPPWPPWRK 12
XX | | | | |
XX Db 1 ILRWPPWPPWRK 12
XX | | | | |
XX RESULT 8
XX ADC98990
XX ID ADC98990 standard; peptide; 14 AA.

```

ADC98989
 ID ADC98989 standard; peptide; 14 AA.
 AC ADC98989;
 XX
 XX
 DT 01-JAN-2004 (first entry)
 DE Synthetic indolicidin analogue peptide - 11D21 CN.
 KW indolicidin analogue; antiseborrheic; dermatological; antiinflammatory;
 KW antiarthritic; immunosuppressive; vulnary; antipruritic; antimicrobial;
 KW antipruritic; neuroprotective; antipsoriatic; inflammation; acne;
 KW arthritis; autoimmune disease; burn; Crohn's; colitis;
 KW contact hypersensitivity; delayed; eczema; endotoxin shock syndrome;
 KW fibromyositis; graft rejection; microbial infection; multiple sclerosis;
 KW parapsoriasis; psoriasis; sclerosis; seborrhea.
 XX
 OS Synthetic.
 XX
 XX
 PH Key Location/Qualifiers
 FT Modified-site 14
 FT /note= "C-terminal amide"
 XX
 XX WO2003018619-A2.
 PN
 XX
 PD 06-MAR-2003.
 XX
 XX 26-AUG-2002; 2002WO-CA001351.
 PF
 XX 24-AUG-2001; 2001US-0315003P.
 PR
 XX 26-AUG-2002; 2002US-00229368.
 XX
 XX (MICR-) MICROLOGIX BIOTECH INC.
 PA (MCNI/) MCNICOL P J.
 PA (PAWL/) PAWLAK S K.
 PA (RUBI/) RUBINCHIK E.
 PA (CAME/) CAMERON D.
 PA (GUAR/) GUARNA M M.
 XX
 PI Mcnicol PJ, Pawlak SK, Rubinchik E, Cameron D, Guarna MM;
 XX
 DR WPI; 2003-393247/37.
 XX
 XX Novel indolicidin analog useful for treating or preventing inflammation
 PT at a target site associated with a condition such as acne, arthritis,
 PT burn, Crohn's disease, colitis, and in image analysis and diagnostic
 PT assays.
 XX
 PS Example 1; Page 56; 66pp; English.
 XX
 XX The invention relates to a novel indolicidin analogue. The analogue of
 CC the invention demonstrates antiseborrheic, dermatological,
 CC antiinflammatory, antiarthritic, immunosuppressive, vulnary,
 CC antipruritic, antimicrobial, antipruritic, neuroprotective and
 CC antipsoriatic and may be useful for treating or preventing inflammation
 CC at a target site. The inflammation at the target site may be associated
 CC with a condition selected from acne, arthritis, autoimmune disease, burn,
 CC Crohn's disease, colitis, contact hypersensitivity, delayed
 CC hypersensitivity, eczema, endotoxin shock syndrome, fibromyositis, graft
 CC rejection, microbial infection, multiple sclerosis, parapsoriasis,
 CC psoriasis, sclerosis and seborrhea. The current sequence is that of the
 XX synthetic indolicidin analogue peptide of the invention.
 XX
 SQ Sequence 14 AA;

Query Match 100.0%; Score 86; DB 7; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.3e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILRWPNPWRK 12
 DB 1 ILRWPNPWRK 12

RESULT 9
 AAY24553
 ID AAY24553 standard; peptide; 20 AA.
 XX
 AC AAY24553;
 XX
 DT 18-AUG-1999 (first entry)
 DE Indolicidin analogue #5.
 XX
 DE Indolicidin; bactericidal infection; photo-oxidised solubiliser;
 KW antimicrobial; antibiotic; antiarrhythmic; surface disinfectant; additive;
 KW shampoo; soap; insecticide; herbicide; preservative; food;
 KW technical material.
 XX
 OS Synthetic.
 XX
 XX WO9807745-A2.
 PN
 XX 26-FEB-1998.
 PD
 XX 21-AUG-1997; 97WO-US014779.
 PF
 XX 21-AUG-1996; 96US-0024754P.
 PR
 XX 13-JAN-1997; 97US-0034949P.
 XX
 XX (MICR-) MICROLOGIX BIOTECH INC.
 PA
 XX Fraser JR, West MH, Krieger TJ, Taylor R, Erfle D;
 PI WPI; 1998-169090/15.
 XX
 DR New indolicidin analogues with antimicrobial activity and related nucleic
 PT acid - vectors, transformed cells and antibodies, also conjugates with
 PT polyoxyethylene glycol and fatty acid to reduce toxicity, useful
 PT therapeutically, as disinfectants etc.
 XX
 PS Claim 11; Page 88; 129pp; English.
 XX
 XX AAY24549 to AAY24615 represent indolicidin analogues of formulae (I)-
 CC (VIII) containing up to 25 amino acids (aa); R₁XX₁XB (I), R₂XX₂XB (II),
 CC R₃XX₃XB (III), R₄XX₄XB (IV), R₅XX₅XB (V), R₆XX₆XB (VI), R₇XX₇XB (VII),
 CC R₈XX₈XB (VIII), R₉XX₉XB (IX), R₁₀XX₁₀XB (X), R₁₁XX₁₁XB (XI), R₁₂XX₁₂XB (XII),
 CC R₁₃XX₁₃XB (XIII), R₁₄XX₁₄XB (XIV), R₁₅XX₁₅XB (XV), R₁₆XX₁₆XB (XVI), R₁₇XX₁₇XB (XVII),
 CC R₁₈XX₁₈XB (XVIII), R₁₉XX₁₉XB (XIX), R₂₀XX₂₀XB (XX), R₂₁XX₂₁XB (XXI),
 CC R₂₂XX₂₂XB (XXII), R₂₃XX₂₃XB (XXIII), R₂₄XX₂₄XB (XXIV), R₂₅XX₂₅XB (XXV). Where Z =
 CC P or V; X = hydrophobic residue, preferably W; B = basic aa, preferably R
 CC or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V; in (VIII) at
 CC least 2 X = F or Y. The analogues are used to treat infections caused by
 CC bacteria (Gram positive or negative, or anaerobic); fungi (yeast or
 CC moulds); parasites (protozoa, nematodes, cestodes or trematodes) or
 CC viruses. Typical of very many pathogens that can be controlled are
 CC Leishmania, Trypanosoma, Ascaris lumbricoide, Fasciola hepatica,
 CC Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus aureus,
 CC Listeria, Clostridium, rotavirus and papilloma virus. Compounds derived
 CC from the analogues may be used similarly; the compounds may also be
 CC prepared from antibiotics or antiarrhythmic agents. The analogues may be
 CC used therapeutically or to coat medical devices; also they are useful as
 CC surface disinfectants, as additives to shampoo or soaps, as insecticides
 CC or herbicides, or as preservatives for foods and technical materials. The
 CC analogues are administered by injection, lavage, orally or topically.
 CC generally at 0.1-50 mg/kg. These analogues have a broader spectrum of
 CC activity than indolicidin and modification as compounds reduces their
 CC toxicity
 XX
 SQ Sequence 20 AA;

Query Match 100.0%; Score 86; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILRWPNPWRK 12
 DB 1 ILRWPNPWRK 12

XX	Result 10
XX	AAAY91797
XX	AAAY91797 standard; peptide; 20 AA.
XX	AAAY91797;
XX	06-JUN-2000 (first entry)
XX	Amino acid sequence of cationic peptide WBI 11B17CN.
XX	Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
XX	leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
XX	breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
XX	multidrug resistance.
XX	Synthetic.
XX	W09965506-A2.
XX	23-DEC-1999.
XX	14-JUN-1999; 99WO-CA000552.
XX	12-JUN-1998; 98US-00096541.
XX	(MICR-) MICROLOGIX BIOTECH INC.
XX	Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
XX	WPI; 2000-223549/19.
XX	Novel pharmaceutical composition containing optionally activated
XX	polyoxyalkylene-modified cationic peptides, useful for treating tumors.
XX	Disclosure; page 15; 94pp; English.
XX	This sequence represents a cationic peptide amino acid sequence, which
XX	can be used in the pharmaceutical composition of the invention. The
XX	invention relates to a pharmaceutical composition containing at least one
XX	activated polyoxyalkylene (APO)-modified cationic peptide. The
XX	modification of peptides with APO increases their activity against tumour
XX	cells, including those with a multidrug resistant phenotype. The
XX	pharmaceutical composition can be used to treat tumours, specifically
XX	lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
XX	cervix, uterus, skin, prostate, liver and colon
XX	Sequence 20 AA;
XX	Query Match 100.0%; Score 86; DB 3; Length 20;
XX	Best Local Similarity 100.0%; Pred. No. 1.8e-05;
XX	Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 ILRWFWPWRRK 12
DB	1 ILRWFWPWRRK 12
XX	Result 11
XX	ADAA00530
XX	ID ADAA00530 standard; peptide; 20 AA.
XX	ADAA00530;
XX	06-NOV-2003 (first entry)
XX	Antimicrobial cationic peptide 11B17CN.
XX	antimicrobial; cationic; viscosity-increasing agent; solvent; buffer;
XX	antibacterial; virucide; antiinflammatory; fungicide; protozoacide;
XX	parasiticide; vulnary; dermatological; herbicide; insecticide;
XX	infection; systemic infection; sepsis; acne; disinfectant; herbicide;
XX	insecticide; silicone sealant.

KW arthritis; autoimmune disease; burn; Crohn's; colitis;
 KW contact hypersensitivity; delayed; eczema; endotoxin shock syndrome;
 KW fibromyositis; graft rejection; microbial infection; multiple sclerosis;
 KW parapsoriasis; psoriasis; sclerosis; seborrhea.
 XX Synthetic.
 OS
 XX WO2003018619-A2.
 PN
 XX
 XX
 XX PD 06-MAR-2003.
 XX
 XX PF 26-AUG-2002; 2002WO-CA001351.
 XX
 XX PR 24-AUG-2001; 2001US-0315003P.
 PR 26-AUG-2002; 2002US-00229368.
 XX
 XX PA (MICR-) MICROLOGIX BIOTECH INC.
 PA (MCNI/) MCNICOL P J.
 PA (PAWL/) PAWLAK S K.
 PA (RUBI/) RUBINCHIK E.
 PA (CAWE/) CAMERON D.
 PA (GUAR/) GUARNA M M.
 XX
 XX MCnicol PJ, Pawlak SK, Rubinchik E, Cameron D, Guarna MM;
 PI
 XX WPI; 2003-393247/37.
 DR
 XX
 XX Novel indolicidin analog useful for treating or preventing inflammation
 PT at a target site associated with a condition such as acne, arthritis,
 PT burn, Crohn's disease, colitis, and in image analysis and diagnostic
 PT assays.
 XX
 XX Example 1; Page 47; 66pp; English.
 PS
 XX The invention relates to a novel indolicidin analogue. The analogue of
 CC the invention demonstrates antiseborrheic, dermatological,
 CC antiinflammatory, antiarthritic, immunosuppressive, vulnerary,
 CC antipruritic, antimicrobial, antipruritic, neuroprotective and
 CC antipsoriatic and may be useful for treating or preventing inflammation
 CC at a target site. The inflammation at the target site may be associated
 CC with a condition selected from acne, arthritis, autoimmune disease, burn,
 CC Crohn's disease, colitis, contact hypersensitivity, delayed
 CC hypersensitivity, eczema, endotoxin shock syndrome, fibromyositis, graft
 CC rejection, microbial infection, multiple sclerosis, parapsoriasis,
 CC psoriasis, sclerosis and seborrhea. The current sequence is that of the
 CC synthetic indolicidin analogue peptide of the invention.
 XX
 XX Sequence 20 AA;
 SQ
 Query Match 100.0%; Score 86; DB 7; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILRWPPWPWRK 12
 |||||
 DB 1 ILRWPPWPWRK 12
 |||||
 RESULT 13
 ID AAW66376
 XX AAW66376 standard; peptide; 21 AA.
 XX
 XX AAW66376;
 AC
 XX
 XX 12-JAN-1999 (first entry)
 DT
 XX
 XX Cationic peptide of claim 15 #3.
 DE
 XX Indolicidin analogue; resistance; cationic peptide; antibiotic;
 KW bacterial infection; tolerance; antibacterial; microorganism; bacteria;
 KW fungus; parasite; virus.
 XX
 XX Synthetic.

XX WO9840401-A2.
 PN
 XX 17-SEP-1998.
 PD
 XX
 XX 10-MAR-1998; 98WO-CA000190.
 XX
 XX 10-MAR-1997; 97US-0040649P.
 PR 20-AUG-1997; 97US-00915314.
 PR 26-SEP-1997; 97US-0060099P.
 PR 25-FEB-1998; 98US-00030619.
 XX
 XX (MICR-) MICROLOGIX BIOTECH INC.
 PA
 XX Fraser JR, West MHP, Mcnicol PJ;
 PI
 XX WPI; 1998-520800/44.
 DR
 XX
 XX New indolicidin peptide analogues - useful for, e.g. enhancing activity
 PT of antibiotic or overcoming tolerance, acquired resistance or inherent
 PT resistance of microorganisms.
 PT
 XX Claim 15; Page 93; 105pp; English.
 PS
 XX The present sequence represents a specifically claimed cationic peptide
 CC from the present invention. The present invention describes compositions
 CC and methods for treating infection, especially bacterial infections. The
 CC compositions and methods use cationic peptides in combination with an
 CC antibiotic agent which are then administered to a patient to enhance the
 CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
 CC acquired resistance; and (c) inherent resistance. The combinations of
 CC antibiotics and cationic peptides can provide synergistic activity
 CC against a microorganism that is tolerant, inherently resistant, or has
 CC acquired resistance to an antibiotic agent. They can be used for killing
 CC e.g. bacteria, fungi, parasites and viruses
 XX
 XX Sequence 21 AA;
 SQ
 Query Match 100.0%; Score 86; DB 2; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILRWPPWPWRK 12
 |||||
 DB 1 ILRWPPWPWRK 12
 |||||
 RESULT 14
 AAY24554
 ID AAY24554 standard; peptide; 21 AA.
 XX
 XX AAY24554;
 AC
 XX
 XX 18-AUG-1999 (first entry)
 DT
 XX
 XX Indolicidin analogue #6.
 DE
 XX Indolicidin; bacterial infection; photo-oxidised solubiliser;
 KW antimicrobial; antibiotic; antiarthritic; surface disinfectant; additive;
 KW shampoo; soap; insecticide; herbicide; preservative; food;
 KW technical material.
 XX
 XX Synthetic.
 OS
 XX WO9807745-A2.
 PN
 XX 26-FEB-1998.
 PD
 XX
 XX 21-AUG-1997; 97WO-US014779.
 PF
 XX 21-AUG-1996; 96US-0024754P.
 PR 13-JAN-1997; 97US-0034949P.
 PR
 XX

(MICR-) MICROLOGIX BIOTECH INC.

Fraser JR, West MH, Krieger TJ, Taylor R, Erfle D;

WPI; 1998-169090/15.

New indolicidin analogues with antimicrobial activity and related nucleic acid - vectors, transformed cells and antibodies, also conjugates with polyoxyalkylene glycol and fatty acid to reduce toxicity, useful therapeutically, as disinfectants etc.

Claim 11; Page 88; 129pp; English.

AAV24549 to AAY24615 represent indolicidin analogues of formulae (I) - (VIII) containing up to 25 amino acids (aa): RXZXZXB (II), BXZXZXB (III), BXBZXZXB (IV), BXZXZXB (V), LBNXZXZXB (VI), LKXZXZXB (VII) and BXBZXZXB (VIII). Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa, preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V; in (VIII) at least 2 X = F or Y. The analogues are used to treat infections caused by bacteria (Gram positive or negative, or anaerobic); fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or trematodes) or viruses. Typical of very many pathogens that can be controlled are Leishmania, Trypanosoma, Ascaris lumbricoidea, Fasciola hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds derived from the analogues may be used similarly; the compounds may also be prepared from antibiotics or antiarrhythmic agents. The analogues may be used therapeutically or to coat medical devices; also they are useful as surface disinfectants, as additives to shampoo or soaps, as insecticides or herbicides, or as preservatives for foods and technical materials. The analogues are administered by injection, lavage, orally or topically, generally at 0.1-50 mg/kg. These analogues have a broader spectrum of activity than indolicidin and modification as compounds reduces their toxicity

Sequence 21 AA;

Query Match 100.0%; Score 86; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILRWPMWPRRK 12

Db 1 ILRWPMWPRRK 12

RESULT 15

AAV24552

ID AAY24552 standard; peptide; 21 AA.

AC AAY24552;

DT 18-AUG-1999 (first entry)

DE Indolicidin analogue #4.

Indolicidin; bacterial infection; photo-oxidised solubiliser; antimicrobial; antibiotic; antiarrhythmic; surface disinfectant; additive; shampoo; soap; insecticide; herbicide; preservative; food; technical material.

OS Synthetic.

XX WO9807745-A2.

XX 26-FEB-1998.

XX 21-AUG-1997; 97WO-US014779.

XX 21-AUG-1996; 96US-0024754P.

XX 13-JAN-1997; 97US-0034949P.

XX

(MICR-) MICROLOGIX BIOTECH INC.

Fraser JR, West MH, Krieger TJ, Taylor R, Erfle D;

WPI; 1998-169090/15.

New indolicidin analogues with antimicrobial activity and related nucleic acid - vectors, transformed cells and antibodies, also conjugates with polyoxyalkylene glycol and fatty acid to reduce toxicity, useful therapeutically, as disinfectants etc.

Claim 11; Page 88; 129pp; English.

AAV24549 to AAY24615 represent indolicidin analogues of formulae (I) - (VIII) containing up to 25 amino acids (aa): RXZXZXB (II), BXZXZXB (III), BXBZXZXB (IV), BXZXZXB (V), LBNXZXZXB (VI), LKXZXZXB (VII) and BXBZXZXB (VIII). Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa, preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V; in (VIII) at least 2 X = F or Y. The analogues are used to treat infections caused by bacteria (Gram positive or negative, or anaerobic); fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or trematodes) or viruses. Typical of very many pathogens that can be controlled are Leishmania, Trypanosoma, Ascaris lumbricoidea, Fasciola hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds derived from the analogues may be used similarly; the compounds may also be prepared from antibiotics or antiarrhythmic agents. The analogues may be used therapeutically or to coat medical devices; also they are useful as surface disinfectants, as additives to shampoo or soaps, as insecticides or herbicides, or as preservatives for foods and technical materials. The analogues are administered by injection, lavage, orally or topically, generally at 0.1-50 mg/kg. These analogues have a broader spectrum of activity than indolicidin and modification as compounds reduces their toxicity

Sequence 21 AA;

Query Match 100.0%; Score 86; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILRWPMWPRRK 12

Db 1 ILRWPMWPRRK 12

Search completed: May 4, 2004, 15:19:39

Job time : 45.7895 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:17:07 ; Search time 13.2632 Seconds
(without alignments)
46.709 Million cell updates/sec

Title: US-09-444-281-36
Perfect score: 86
Sequence: 1 ILRWPWPWRRK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	100.0	12	3	US-08-915-314-42
2	86	100.0	12	4	US-09-030-619-23
3	86	100.0	12	4	US-09-667-486-42
4	86	100.0	20	3	US-08-915-314-47
5	86	100.0	20	4	US-09-030-619-24
6	86	100.0	20	4	US-09-667-486-47
7	86	100.0	21	3	US-08-915-314-46
8	86	100.0	21	3	US-08-915-314-48
9	86	100.0	21	4	US-09-030-619-47
10	86	100.0	21	4	US-09-030-619-48
11	86	100.0	21	4	US-09-667-486-46
12	86	100.0	21	4	US-09-667-486-48
13	86	100.0	28	4	US-09-030-619-50
14	86	100.0	28	4	US-09-030-619-104
15	83	95.5	12	3	US-08-915-314-40
16	83	95.5	12	4	US-09-030-619-43
17	83	95.5	12	4	US-09-667-486-40
18	82	95.3	12	3	US-08-915-314-76
19	82	95.3	12	4	US-09-030-619-30
20	82	95.3	12	4	US-09-030-619-111
21	82	95.3	12	4	US-09-667-486-76
22	81	94.2	12	3	US-08-915-314-77
23	81	94.2	12	3	US-08-915-314-87
24	81	94.2	12	4	US-09-030-619-82
25	81	94.2	12	4	US-09-030-619-91
26	81	94.2	12	4	US-09-667-486-71
27	81	94.2	12	4	US-09-667-486-87

28	80	93.0	12	3	US-08-915-314-78	Sequence 78, Appl
29	80	93.0	12	3	US-08-915-314-85	Sequence 85, Appl
30	80	93.0	12	3	US-08-915-314-86	Sequence 86, Appl
31	80	93.0	12	4	US-09-030-619-83	Sequence 83, Appl
32	80	93.0	12	4	US-09-030-619-89	Sequence 89, Appl
33	80	93.0	12	4	US-09-030-619-90	Sequence 90, Appl
34	80	93.0	12	4	US-09-667-486-78	Sequence 78, Appl
35	80	93.0	12	4	US-09-667-486-85	Sequence 85, Appl
36	80	93.0	12	4	US-09-667-486-86	Sequence 86, Appl
37	78	90.7	12	3	US-08-915-314-80	Sequence 80, Appl
38	78	90.7	12	3	US-08-915-314-83	Sequence 83, Appl
39	78	90.7	12	4	US-09-030-619-31	Sequence 31, Appl
40	78	90.7	12	4	US-09-030-619-40	Sequence 40, Appl
41	78	90.7	12	4	US-09-030-619-87	Sequence 87, Appl
42	78	90.7	12	4	US-09-667-486-80	Sequence 80, Appl
43	78	90.7	12	4	US-09-667-486-83	Sequence 83, Appl
44	78	90.7	13	3	US-08-915-314-38	Sequence 38, Appl
45	78	90.7	13	4	US-09-030-619-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1
US-08-915-314-42
; Sequence 42, Application US/08915314
; Patent No. 6180604
; GENERAL INFORMATION:
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,314
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6180604tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 660081.405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-915-314-42

Query Match 100.0%; Score 86; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ILRWPWPWRRK 12
|||||

Db 1 ILRWPWPWRK 12

RESULT 5

US-09-030-619-24

; Sequence 24, Application US/09030619B

; Patent No. 6503881

; GENERAL INFORMATION:

; APPLICANT: Krieger, Robert

; APPLICANT: Taylor, Robert

; APPLICANT: Erfile, Douglas

; APPLICANT: Fraser, Janet R.

; APPLICANT: West, Michael H.P.

; APPLICANT: McNicol, Patricia J.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION

; TITLE OF INVENTION: WITH ANTIBIOTICS

; FILE REFERENCE: 660081.406

; CURRENT APPLICATION NUMBER: US/09/030,619B

; CURRENT FILING DATE: 1998-02-25

; NUMBER OF SEQ ID NOS: 232

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 24

; LENGTH: 20

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Indolicidin Analogue

US-09-030-619-24

Query Match

Best Local Similarity 100.0%; Score 86; DB 4; Length 20;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRWPWPWRK 12

|||||

Db 1 ILRWPWPWRK 12

RESULT 6

US-09-667-486-47

; Sequence 47, Application US/09667486

; Patent No. 6538106

; GENERAL INFORMATION:

; APPLICANT: Fraser, Janet R.

; APPLICANT: West, Michael H.P.

; APPLICANT: Krieger, Timothy J.

; APPLICANT: Taylor, Robert

; APPLICANT: Erfile, Douglas

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

; TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN

; NUMBER OF SEQUENCES: 90

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/667,486

; FILING DATE: 22-Sep-2000

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/915,314

; FILING DATE: 20-AUG-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 6538106tenburg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 660081.405

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 47:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 47:

US-09-667-486-47

Query Match

Best Local Similarity 100.0%; Score 86; DB 4; Length 20;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRWPWPWRK 12

|||||

Db 1 ILRWPWPWRK 12

RESULT 7

US-08-915-314-46

; Sequence 46, Application US/08915314

; Patent No. 6180604

; GENERAL INFORMATION:

; APPLICANT: Fraser, Janet R.

; APPLICANT: West, Michael H.P.

; APPLICANT: Krieger, Timothy J.

; APPLICANT: Taylor, Robert

; APPLICANT: Erfile, Douglas

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

; TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN

; NUMBER OF SEQUENCES: 90

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/915,314

; FILING DATE: 20-AUG-1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 6180604tenburg Ph.D., Carol

; REGISTRATION NUMBER: 39,317

; REFERENCE/DOCKET NUMBER: 660081.405

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:

LENGTH: 21 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-915-314-46

Query Match

Best Local Similarity 100.0%; Score 86; DB 3; Length 21;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRWPWPWRK 12

|||||

```
Db 1 ILRWPWPWRK 12
;
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 47
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
;
US-09-030-619-47
Query Match 100.0%; Score 86; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRWPWPWRK 12
| | | | | | | | | |
Db 1 ILRWPWPWRK 12
| | | | | | | | | |

RESULT 10
US-09-030-619-48
; Sequence 48, Application US/09030619B
; Patent No. 6503881
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfle, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
;
US-09-030-619-48
Query Match 100.0%; Score 86; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRWPWPWRK 12
| | | | | | | | | |
Db 1 ILRWPWPWRK 12
| | | | | | | | | |

RESULT 11
US-09-667-486-46
; Sequence 46, Application US/09667486
; Patent No. 6538106
; GENERAL INFORMATION:
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfle, Douglas
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
;

Db 1 ILRWPWPWRK 12
| | | | | | | | | |
Db 1 ILRWPWPWRK 12
| | | | | | | | | |

RESULT 9
US-09-030-619-47
; Sequence 47, Application US/09030619B
; Patent No. 6503881
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfle, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
;

US-08-915-314-48
; Sequence 48, Application US/08915314
; Patent No. 6180604
; GENERAL INFORMATION:
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfle, Douglas
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,314
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6180604tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 660081.405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-915-314-48
Query Match 100.0%; Score 86; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRWPWPWRK 12
| | | | | | | | | |
Db 1 ILRWPWPWRK 12
| | | | | | | | | |
```

/ COUNTRY: USA
/ ZIP: 98104
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/667,486
/ FILING DATE: 22-Sep-2000
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/915,314
/ FILING DATE: 20-AUG-1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: No. 6538106tenburg Ph.D., Carol
/ REGISTRATION NUMBER: 39,317
/ REFERENCE/DOCKET NUMBER: 660081.405
/ TELEPHONE: (206) 622-4900
/ TELEFAX: (206) 682-6031
/ INFORMATION FOR SEQ ID NO: 46:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 21 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-667-486-46

Query Match 100.0%; Score 86; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRWPWPWRRK 12
| | | | | | | | | | | | | | | | | | | | |
Db 1 ILRWPWPWRRK 12
| | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-09-667-486-48
; Sequence 48, Application US/09667486
; Patent No. 6538106
; GENERAL INFORMATION:
; APPLICANT: Fraser, Janet R.
; Applicant: West, Michael H.P.
; Applicant: Krieger, Timothy J.
; Applicant: Erile, Douglas
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/667,486
FILING DATE: 22-Sep-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: No. 6538106tenburg Ph.D., Carol

/ REGISTRATION NUMBER: 39,317
/ REFERENCE/DOCKET NUMBER: 660081.405
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (206) 622-4900
/ TELEFAX: (206) 682-6031
/ INFORMATION FOR SEQ ID NO: 48:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 21 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-09-667-486-48

Query Match 100.0%; Score 86; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRWPWPWRRK 12
| | | | | | | | | | | | | | | | | | | | |
Db 1 ILRWPWPWRRK 12
| | | | | | | | | | | | | | | | | | | | |

RESULT 13
US-09-030-619-50
; Sequence 50, Application US/09030619B
; Patent No. 6503881
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; Applicant: Taylor, Robert
; Applicant: Erile, Douglas
; Applicant: Fraser, Janet R.
; Applicant: West, Michael H.P.
; Applicant: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 50
LENGTH: 28
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-50

Query Match 100.0%; Score 86; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 6.1e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRWPWPWRRK 12
| | | | | | | | | | | | | | | | | | | | |
Db 1 ILRWPWPWRRK 12
| | | | | | | | | | | | | | | | | | | | |

RESULT 14
US-09-030-619-104
; Sequence 104, Application US/09030619B
; Patent No. 6503881
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; Applicant: Taylor, Robert
; Applicant: Erile, Douglas
; Applicant: Fraser, Janet R.
; Applicant: West, Michael H.P.
; Applicant: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 50
LENGTH: 28
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-104

FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030.619H
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 104
LENGTH: 28
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-104

Search completed: May 4, 2004, 15:23:52
Job time : 13.2632 secs

Query Match 100.0%; Score 86; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 6.1e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRWPWPWPWRK 12

Db 1 ILRWPWPWPWRK 12

RESULT 15
US-08-915-314-40
Sequence 40, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

US-08-915-314-40
Query Match 96.5%; Score 83; DB 3; Length 12;
Best Local Similarity 91.7%; Pred. No. 6.8e-06;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRWPWPWPWRK 12

Db 1 ILKWPWPWPWRK 12

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:22:18 ; Search time 34.4211 Seconds
(without alignments)
96.635 Million cell updates/sec

Title: US-09-444-281-36
Perfect score: 86
Sequence: 1 ILRWPPWPPRRK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	86	100.0	12	9 US-09-030-619-23
2	86	100.0	12	12 US-10-277-232-23
3	86	100.0	12	14 US-10-229-368-23
4	86	100.0	12	14 US-10-229-368-55
5	86	100.0	12	14 US-10-225-087-23
6	86	100.0	12	14 US-10-225-087-119
7	86	100.0	12	15 US-10-395-896-12
8	86	100.0	12	15 US-10-395-896-60
9	86	100.0	12	15 US-10-395-896-68
10	86	100.0	12	15 US-10-277-232-23
11	86	100.0	12	15 US-10-351-985-42
12	86	100.0	20	9 US-09-030-619-24
13	86	100.0	20	12 US-10-277-232-24
14	86	100.0	20	14 US-10-229-368-30
15	86	100.0	20	14 US-10-229-368-30

16	86	100.0	20	14 US-10-225-087-29	Sequence 29, Appl
17	86	100.0	20	15 US-10-277-232-24	Sequence 24, Appl
18	86	100.0	20	15 US-10-351-985-47	Sequence 47, Appl
19	86	100.0	21	9 US-09-030-619-47	Sequence 47, Appl
20	86	100.0	21	9 US-09-030-619-48	Sequence 48, Appl
21	86	100.0	21	12 US-10-277-232-47	Sequence 47, Appl
22	86	100.0	21	12 US-10-277-232-48	Sequence 48, Appl
23	86	100.0	21	14 US-10-229-368-29	Sequence 29, Appl
24	86	100.0	21	14 US-10-229-368-31	Sequence 31, Appl
25	86	100.0	21	14 US-10-225-087-28	Sequence 28, Appl
26	86	100.0	21	14 US-10-225-087-30	Sequence 30, Appl
27	86	100.0	21	15 US-10-277-233-47	Sequence 47, Appl
28	86	100.0	21	15 US-10-277-233-48	Sequence 48, Appl
29	86	100.0	21	15 US-10-351-985-46	Sequence 46, Appl
30	86	100.0	21	15 US-10-351-985-48	Sequence 48, Appl
31	86	100.0	24	15 US-10-395-896-47	Sequence 47, Appl
32	86	100.0	24	15 US-10-395-896-48	Sequence 48, Appl
33	86	100.0	25	15 US-10-395-896-49	Sequence 49, Appl
34	86	100.0	25	15 US-10-395-896-50	Sequence 50, Appl
35	86	100.0	25	15 US-10-395-896-51	Sequence 51, Appl
36	86	100.0	25	15 US-10-395-896-52	Sequence 52, Appl
37	86	100.0	25	15 US-10-395-896-53	Sequence 53, Appl
38	86	100.0	26	15 US-10-395-896-54	Sequence 54, Appl
39	86	100.0	26	15 US-10-395-896-55	Sequence 55, Appl
40	86	100.0	26	15 US-10-395-896-56	Sequence 56, Appl
41	86	100.0	26	15 US-10-395-896-57	Sequence 57, Appl
42	86	100.0	28	9 US-09-030-619-50	Sequence 50, Appl
43	86	100.0	28	9 US-09-030-619-104	Sequence 104, Appl
44	86	100.0	28	12 US-10-277-232-50	Sequence 50, Appl
45	86	100.0	28	12 US-10-277-232-104	Sequence 104, Appl

ALIGNMENTS

RESULT 1

US-09-030-619-23
; Sequence 23, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 66081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-23

Query Match 100.0%; Score 86; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00057; Gaps 0;
Matches 12; Conservative 0; Mismatches 0; Indels 0;

QY 1 ILRWPPWPPRRK 12
|||||
Db 1 ILRWPPWPPRRK 12

RESULT 2

US-10-277-232-23

```
; Sequence 23, Application US/10277232
; Publication No. US20030211537A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Brife, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.418
; CURRENT APPLICATION NUMBER: US/10/277,232
; CURRENT FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-10-277-232-23

Query Match          100.0%; Score 86; DB 12; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00057; Mismatches 0; Indels 0; Gaps 0;
Matches 12; Conservative 0;

Qy 1 ILRWPWPWRRK 12
Db 1 ILRWPWPWRRK 12
|||||

RESULT 3
US-10-229-368-23
; Sequence 23, Application US/10229368
; Publication No. US20030148945A1
; GENERAL INFORMATION:
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Pawlak, Sonia K.
; APPLICANT: Rubinchik, Evelina
; APPLICANT: Cameron, Dale
; APPLICANT: Guarna, Maria Marta
; TITLE OF INVENTION: ANTIMICROBIAL AND ANTI-INFLAMMATORY
; TITLE OF INVENTION: PEPTIDES
; FILE REFERENCE: 660081.418
; CURRENT APPLICATION NUMBER: US/10/229,368
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin peptide analogs
US-10-229-368-23

Query Match          100.0%; Score 86; DB 14; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00057; Mismatches 0; Indels 0; Gaps 0;
Matches 12; Conservative 0;

Qy 1 ILRWPWPWRRK 12
Db 1 ILRWPWPWRRK 12
|||||

RESULT 4
US-10-229-368-55
; Sequence 55, Application US/10229368
; Publication No. US20030148945A1
; GENERAL INFORMATION:
```

```
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Pawlak, Sonia K.
; APPLICANT: Rubinchik, Evelina
; APPLICANT: Cameron, Dale
; APPLICANT: Guarna, Maria Marta
; TITLE OF INVENTION: ANTIMICROBIAL AND ANTI-INFLAMMATORY
; TITLE OF INVENTION: PEPTIDES
; FILE REFERENCE: 660081.418
; CURRENT APPLICATION NUMBER: US/10/229,368
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin peptide analogs
US-10-229-368-55

Query Match          100.0%; Score 86; DB 14; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00057; Mismatches 0; Indels 0; Gaps 0;
Matches 12; Conservative 0;

Qy 1 ILRWPWPWRRK 12
Db 1 ILRWPWPWRRK 12
|||||

RESULT 5
US-10-225-087-23
; Sequence 23, Application US/10225087
; Publication No. US20030171281A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Fraser, Janet R.
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES AND
; TITLE OF INVENTION: FORMULATIONS THEREOF
; FILE REFERENCE: 660081.417
; CURRENT APPLICATION NUMBER: US/10/225,087
; CURRENT FILING DATE: 2003-01-10
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin analog
US-10-225-087-23

Query Match          100.0%; Score 86; DB 14; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00057; Mismatches 0; Indels 0; Gaps 0;
Matches 12; Conservative 0;

Qy 1 ILRWPWPWRRK 12
Db 1 ILRWPWPWRRK 12
|||||

RESULT 6
US-10-225-087-119
; Sequence 119, Application US/10225087
; Publication No. US20030171281A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Fraser, Janet R.
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES AND
; TITLE OF INVENTION: FORMULATIONS THEREOF
; FILE REFERENCE: 660081.417
; CURRENT APPLICATION NUMBER: US/10/225,087
```

; CURRENT FILING DATE: 2003-01-10
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 119
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin analog
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)-(1)
; OTHER INFORMATION: N - terminal modification (acryloyl)
US-10-225-087-119

Query Match 100.0%; Score 86; DB 14; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILRWPPWPWRRK 12
| | | | | | | | | | | | | |
DB 1 ILRWPPWPWRRK 12

RESULT 7

US-10-395-896-12
; Sequence 12, Application US/10395896
; Publication No. US20030219854A1
; GENERAL INFORMATION:
; APPLICANT: Guarna, Maria Marta
; APPLICANT: Chen, Yuchen
; APPLICANT: Cory, Robert
; APPLICANT: Brinkman, Jacqui
; APPLICANT: Cabralda, Jennifer
; APPLICANT: Metlitskaia, Luba
; APPLICANT: Suleman, Dinar
; TITLE OF INVENTION: METHODS FOR PRODUCING MODIFIED
; FILE REFERENCE: 660081.421
; CURRENT APPLICATION NUMBER: US/10/395,896
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin analogue 11B7
US-10-395-896-12

Query Match 100.0%; Score 86; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILRWPPWPWRRK 12
| | | | | | | | | | | | | |
DB 1 ILRWPPWPWRRK 12

RESULT 8

US-10-395-896-60
; Sequence 60, Application US/10395896
; Publication No. US20030219854A1
; GENERAL INFORMATION:
; APPLICANT: Guarna, Maria Marta
; APPLICANT: Chen, Yuchen
; APPLICANT: Cory, Robert
; APPLICANT: Brinkman, Jacqui
; APPLICANT: Cabralda, Jennifer
; APPLICANT: Metlitskaia, Luba
; APPLICANT: Suleman, Dinar
; TITLE OF INVENTION: METHODS FOR PRODUCING MODIFIED

; TITLE OF INVENTION: ANTI-INFECTIVE PEPTIDES
; FILE REFERENCE: 660081.421
; CURRENT APPLICATION NUMBER: US/10/395,896
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin analogue 11B7CN
; FEATURE:
; NAME/KEY: AMIDATION
; LOCATION: 12
US-10-395-896-60

Query Match 100.0%; Score 86; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILRWPPWPWRRK 12
| | | | | | | | | | | | | |
DB 1 ILRWPPWPWRRK 12

RESULT 9

US-10-395-896-68
; Sequence 68, Application US/10395896
; Publication No. US20030219854A1
; GENERAL INFORMATION:
; APPLICANT: Guarna, Maria Marta
; APPLICANT: Chen, Yuchen
; APPLICANT: Cory, Robert
; APPLICANT: Brinkman, Jacqui
; APPLICANT: Cabralda, Jennifer
; APPLICANT: Metlitskaia, Luba
; APPLICANT: Suleman, Dinar
; TITLE OF INVENTION: METHODS FOR PRODUCING MODIFIED
; FILE REFERENCE: 660081.421
; CURRENT APPLICATION NUMBER: US/10/395,896
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide formed during process for amidation of
; OTHER INFORMATION: 11B7
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: 1
; OTHER INFORMATION: Di-tert-butyl dicarboxylate
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: 11
; OTHER INFORMATION: Di-tert-butyl dicarboxylate
US-10-395-896-68

Query Match 100.0%; Score 86; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILRWPPWPWRRK 12
| | | | | | | | | | | | | |
DB 1 ILRWPPWPWRRK 12

RESULT 10

US-10-395-896-69

Sequence 69, Application US/10395896
Publication No. US20030219854A1
GENERAL INFORMATION:
APPLICANT: Guarna, Maria Marta
APPLICANT: Chen, Yuchen
APPLICANT: Cory, Robert
APPLICANT: Brinkman, Jacqui
APPLICANT: Cabralda, Jennifer
APPLICANT: Metlitskaia, Luba
APPLICANT: Suleman, Dinar
TITLE OF INVENTION: METHODS FOR PRODUCING MODIFIED
TITLE OF INVENTION: ANTI-INFECTION PEPTIDES
FILE REFERENCE: 660081.421
CURRENT APPLICATION NUMBER: US/10/395,896
CURRENT FILING DATE: 2003-03-21
NUMBER OF SEQ ID NOS: 70
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 69
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide formed during process for amidation of
OTHER INFORMATION: 11B7
FEATURE:
NAME/KEY: MOD_RES
LOCATION: 1
OTHER INFORMATION: Di-tert-butyl dicarbonate
FEATURE:
NAME/KEY: MOD_RES
LOCATION: 11
OTHER INFORMATION: Di-tert-butyl dicarbonate and amidation
US-10-395-896-69
Query Match 100.0%; Score 86; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ILRWPWPWRRK 12
Db 1 ILRWPWPWRRK 12
RESULT 11
US-10-277-233-23
Sequence 23, Application US/10277233
Publication No. US20030232750A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTI-BIOTICS
FILE REFERENCE: 660081.406C1
CURRENT APPLICATION NUMBER: US/10/277,233
CURRENT FILING DATE: 2002-10-18
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indolicidin Analogue
US-10-277-233-23
Query Match 100.0%; Score 86; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRWPWPWRRK 12
Db 1 ILRWPWPWRRK 12
RESULT 12
US-10-351-985-42
Sequence 42, Application US/10351985
Publication No. US20040009910A1
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Erfle, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed IP Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/351,985
FILING DATE: 24-Jan-2003
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Pepe, Jeff C.
REGISTRATION NUMBER: 46,985
REFERENCE/DOCKET NUMBER: 660081.405C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-10-351-985-42
Query Match 100.0%; Score 86; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ILRWPWPWRRK 12
Db 1 ILRWPWPWRRK 12
RESULT 13
US-09-030-619-24
Sequence 24, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION

; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-24

Query Match 100.0%; Score 86; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00085;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILRWPPWPWRK 12
|||||
DB 1 ILRWPPWPWRK 12
|||||

RESULT 14
US-10-277-232-24
; Sequence 24, Application US/10277232
; Publication No. US20030211537A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406C1
; CURRENT FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-10-277-232-24

Query Match 100.0%; Score 86; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00085;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILRWPPWPWRK 12
|||||
DB 1 ILRWPPWPWRK 12
|||||

RESULT 15
US-10-229-368-30
; Sequence 30, Application US/10229368
; Publication No. US20030148945A1
; GENERAL INFORMATION:
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Pawlak, Sonia K.
; APPLICANT: Rubinchik, Evelina
; APPLICANT: Cameron, Dale
; APPLICANT: Guarna, Maria Marta
; TITLE OF INVENTION: ANTIMICROBIAL AND ANTI-INFLAMMATORY
; TITLE OF INVENTION: PEPTIDES
; FILE REFERENCE: 660081.418
; CURRENT APPLICATION NUMBER: US/10/229,368

; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin peptide analogs
US-10-229-368-30

Query Match 100.0%; Score 86; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00085;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILRWPPWPWRK 12
|||||
DB 1 ILRWPPWPWRK 12
|||||

Search completed: May 4, 2004, 15:35:36
Job time : 35.4211 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:15:37 ; Search time 11.0526 Seconds
(without alignments)
104.437 Million cell updates/sec

Title: US-09-444-281-36

Perfect score: 86

Sequence: 1 ILRWPWPWRRK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*
1: Pirl.*
2: Pirl2.*
3: Pirl3.*
4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	81.4	144	JC1222	indolicidin precu
2	53	61.6	1173	1 VGIHHC	E2 glycoprotein pr
3	51	59.3	299	2 T12505	hypothetical prote
4	50	58.1	111	2 T29295	hypothetical prote
5	49	57.0	467	2 E89605	protein F1865.2 [1
6	49	57.0	498	1 JT0751	ferredoxin-NADP re
7	48.5	56.4	114	2 T36208	hypothetical prote
8	48	55.8	265	2 AH0755	conserved hypothet
9	47	54.7	248	2 S23449	NADH oxidase (H2O2
10	47	54.7	253	2 G70715	hypothetical prote
11	47	54.7	276	2 B83161	probable short-cha
12	47	54.7	715	2 B70741	probable moey prot
13	47	54.7	1411	2 T48529	hypothetical prote
14	46	53.5	728	2 T51071	related to trfa pr
15	45.5	52.9	505	2 A39128	anthranilate synth
16	45	52.3	187	2 AC3353	hypothetical prote
17	45	52.3	196	2 S55483	modulator of drug
18	45	52.3	273	2 F82646	monofunctional bio
19	45	52.3	412	2 A83604	probable MFS trans
20	45	52.3	448	2 H72376	hypothetical prote
21	45	52.3	1108	2 A48508	cyclic-nucleotide
22	44	51.2	67	2 AC1954	hypothetical prote
23	44	51.2	257	2 S70177	yfjg protein - Yer
24	44	51.2	353	2 AB1823	hypothetical prote
25	44	51.2	361	2 A36669	3-alpha-galactosyl
26	44	51.2	397	2 B70763	probable membrane
27	44	51.2	535	2 T38244	hypothetical prote
28	44	51.2	621	2 S37664	peplomeric polypro
29	44	51.2	630	2 S37663	peplomeric polypro

RESULT 1

JC1222
indolicidin precursor - bovine
N:Alternate names: antimicrobial peptide
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JC1222; A42387; S25664
R:del Sar, G.; Storici, P.; Schneider, C.; Romeo, D.; Zanetti, M.
Biochem. Biophys. Res. Commun. 187, 467-472, 1992
A:Title: CDNA cloning of the neutrophil bactericidal peptide indolicidin.
A:Reference number: JC1222; MUID:92392368; PMID:1520337
A:Accession: JC1222
A:Molecule type: mRNA
A:Residues: 1-144 <SEL>
A:Cross-references: EMBL:X67340; NID:9462; PIDN:CAA47755.1; PID:9463
R:Seibel, M.E.; Novotny, M.J.; Morris, W.L.; Tang, Y.Q.; Smith, W.; Cullor, J.S.
J. Biol. Chem. 267, 4292-4295, 1992
A:Title: Indolicidin, a novel bactericidal tridecapeptide amide from neutrophils.
A:Reference number: A42387; MUID:92165771; PMID:1537821
A:Accession: A42387

ALIGNMENTS

30	44	51.2	967	2	C70831	probable mmpL4 pro
31	44	51.2	968	2	F70746	probable mmpL2 pro
32	44	51.2	968	2	T00322	hypothetical prote
33	44	51.2	1154	1	VGIHIB	E2 glycoprotein pr
34	44	51.2	1162	1	VGIHAX	E2 glycoprotein pr
35	44	51.2	1162	2	S07421	E2 glycoprotein pr
36	44	51.2	1162	2	S14939	E2 glycoprotein pr
37	44	51.2	1162	2	S14940	E2 glycoprotein pr
38	43.5	50.6	276	2	AH0244	probable esterase
39	43.5	50.6	1529	2	A59189	ATP-binding cassette
40	43	50.0	51	2	S23291	light-harvesting p
41	43	50.0	192	2	H86543	hypothetical prote
42	43	50.0	192	2	D72081	conserved hypothet
43	43	50.0	236	2	J00606	arylesterase (EC 3
44	43	50.0	250	2	A83506	probable cobalamin
45	43	50.0	278	2	T46458	hypothetical prote

Query Match 81.4%; Score 70; DB 1; Length 144;

Best Local Similarity 88.9%; Fred. No. 0.014; Mismatches 1; Indels 0; Gaps 0;

Matches 8; Conservative 1; Indels 0; Gaps 0;
QY 3 RWPWPWRR 11
DB 135 KWPWPWRR 143

RESULT 2

VGIHHC
E2 glycoprotein precursor - human coronavirus (strain 229E)
N:Alternate names: peplomer glycoprotein; S glycoprotein; spike glycoprotein
C:Species: human coronavirus
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000
C:Accession: A34766; S05460
R:Raabe, T.; Schelle-Prinz, B.; Siddell, S.G.
J. Gen. Virol. 71, 1065-1073, 1990

A;Title: Nucleotide sequence of the gene encoding the spike glycoprotein of human coronavirus 229E
A;Reference number: A34766; MUID:90264837; PMID:2345367
A;Accession: A34766
A;Molecule type: mRNA
A;Residues: 1-1173 <RAA>
A;Cross-references: EMBL:X16816; NID:G58926; PIDN:CAA34723.1; PID:G58927
A;Experimental source: strain 229E
R;Raabe, T.; Siddell, S.
Nucleic Acids Res. 17, 6387, 1989
A;Title: Nucleotide sequence of the human coronavirus HCV 229E mRNA 4 and mRNA 5 unique
A;Reference number: A34038; MUID:89366667; PMID:2701946
A;Accession: S05460
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1159-1173 <RA2>
A;Cross-references: EMBL:X15654; NID:G58921; PIDN:CAA33680.1; PID:G334827
C;Superfamily: coronavirus E2 glycoprotein
C;Keywords: glycoprotein; transmembrane protein
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-1173/Product: E2 glycoprotein #status predicted <MAT>
F;116-1198/Domain: transmembrane #status predicted <TMN>
F;23,62,98,147,171,176,220,243,326,333,440,464,518,538,542,568,581,587,663,671,930,1015,
F;23,62,98,147,171,176,220,243,326,333,440,464,518,538,542,568,581,587,663,671,930,1015,
Query Match 61.6%; Score 53; DB 1; Length 1173;
Best Local Similarity 62.5%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 2 LRWPWPWP 9
Db 1112 IKWPWWVW 1119
:::|||||
RESULT 3
T12505
hypothetical protein DKFPz434C192.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 23-Jul-1999
C;Accession: T12505
R;Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z17527
A;Accession: T12505
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-299 <ANS>
A;Cross-references: EMBL:AL096753
A;Experimental source: adult testis; clone DKFPz434C192
C;Genetics:
A;Note: DKFPz434C192.1
Query Match 59.3%; Score 51; DB 2; Length 299;
Best Local Similarity 85.7%; Pred. No. 7.9;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 PWPWPWR 11
Db 37 PWPWPWR 43
:::|||||
RESULT 4
T29295
hypothetical protein C50F7.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T29295
R;Johnson, D.; Stellives, L.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid C50F7.
A;Reference number: Z20601
A;Accession: T29295
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-111 <JOH>

A;Cross-references: EMBL:U41557; PIDN:AAA83303.1; CESP:C50F7.8
C;Genetics:
A;Gene: CESP:C50F7.8
Query Match 58.1%; Score 50; DB 2; Length 111;
Best Local Similarity 54.5%; Pred. No. 4.1;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILRWPWPWR 11
Db 12 VWWPWPWPGR 22
:::|||||
RESULT 5
E89605
protein Flag5.2 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Nov-2001
C;Accession: E89605
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: E89605
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-467 <STO>
A;Cross-references: GB:chr_X; PIDN:AAA81082.1; PID:G1055093; GSPDB:GN000028; CESP:F18G5.
C;Genetics:
A;Gene: F18G5.2
A;Map position: X
Query Match 57.0%; Score 49; DB 2; Length 467;
Best Local Similarity 83.3%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 WPWPWP 9
Db 201 WPWPWP 206
:::|||||
RESULT 6
JT0751
ferredoxin-NADP reductase (EC 1.18.1.2), long form precursor - bovine
N;Alternate names: adrenodoxin reductase
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Jul-1994 #sequence_revision 18-Oct-1996 #text_change 03-Jun-2002
C;Accession: JT0751; JT0079; J50390; S03558; P50003; A29504; S52100
R;Takata, Y.; Sagara, Y.; Kono, A.; Sekimizu, K.; Horiuchi, T.
Biol. Pharm. Bull. 16, 1200-1206, 1993
A;Title: Gene structure of bovine adrenodoxin reductase.
A;Reference number: JT0751; MUID:94177140; PMID:8130767
A;Accession: JT0751
A;Molecule type: DNA
A;Residues: 1-498 <TAK>
A;Cross-references: GB:D83475; NID:G1199916; PIDN:BA11921.1; PID:G4521308
A;Experimental source: adrenal cortex
A;Note: the authors translated the codon GTC for residue 205 as Gly
R;Sagara, Y.; Takata, Y.; Miyata, T.; Hara, T.; Horiuchi, T.
J. Biochem. 102, 1333-1336, 1987
A;Title: Cloning and sequence analysis of adrenodoxin reductase cDNA from bovine adrenal
A;Reference number: JT0079; MUID:86198050; PMID:3448086
A;Accession: JT0079
A;Molecule type: mRNA
A;Residues: 1-204,211-498 <SAG>
A;Cross-references: GB:D00211; NID:G217433; PIDN:BAA00150.1; PID:G217434
A;Note: the deduced sequence is partially confirmed by amino acid sequencing of 15 isol
R;Sagara, Y.
submitted to DBJ, September 1989
A;Reference number: J50390
A;Contents: revision, insertion of residues 205-210

QY 5 PWPWP 9
Db 179 PWPWP 183

RESULT 10
G70715
Hypothetical protein Rv0945 - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: G70715
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
A: Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID: 98295987; PMID: 9634230
A: Accession: G70715
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-253 <COL>
A: Cross-references: GB:Z79700; GB:AL123456; NID:G3261628; PIDN: CAB02005.1; PID: g1524217
A: Experimental source: strain H37Rv
C: Genetics:
A: Gene: Rv0945
C: Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
F: 8-190/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 54.7%; Score 47; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PWPWP 9
Db 230 PWPWP 234

RESULT 11
B83161
Probable short-chain dehydrogenase PA3883 [imported] - Pseudomonas aeruginosa (strain PA
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83161
R: Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
A: Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
Nature 406, 959-964, 2000
A: Reference number: A82950; MUID: 20437337; PMID: 10984043
A: Accession: B83161
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-276 <STO>
A: Cross-references: GB:AE004805; GB:AE004091; NID: g9950055; PIDN: AAG07270.1; GSPDB: GN001
A: Experimental source: strain PA01
C: Genetics:
A: Gene: PA3883
C: Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 54.7%; Score 47; DB 2; Length 276;
Best Local Similarity 70.0%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 RWPWPWRX 12
Db 197 RSPWPPLRQ 206

RESULT 12
B70741
Probable moey protein - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: B70741
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
A: Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID: 98295987; PMID: 9634230
A: Accession: B70741
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-715 <COL>
A: Cross-references: GB:Z75555; GB:AL123456; NID: G3261608; PIDN: CAA99988.1; PID: e2503556;
A: Experimental source: strain H37Rv
C: Genetics:
A: Gene: moey

Query Match 54.7%; Score 47; DB 2; Length 715;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWPWPWRX 11
Db 65 RWAYYPWR 73

RESULT 13
T48529
Hypothetical protein T22P22.90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T48529
R: Bevan, M.; Hilbert, H.; Braun, M.; Holzner, E.; Brandt, A.; Duesterhoeft, A.; Bancroft
submitted to the Protein Sequence Database, April 2000
A: Reference number: Z24490
A: Accession: T48529
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-1411 <BEV>
A: Cross-references: EMBL:AL163814
A: Experimental source: cultivar Columbia; BAC clone T22P22
C: Genetics:
A: Map position: 5
A: Introns: 281/2; 320/1; 389/3; 429/3; 473/3; 515/3; 534/2; 567/3; 602/1; 669/1; 776/2;
A: Note: T22P22.90

Query Match 54.7%; Score 47; DB 2; Length 1411;
Best Local Similarity 63.6%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 LRWPWPWRX 12
Db 1013 LAWSQQWRK 1023

RESULT 14
T51071
Related to trfa protein [imported] - Neurospora crassa
N: Alternate names: protein B2A19.50
C:Species: Neurospora crassa
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T51071
R: Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura
submitted to the Protein Sequence Database, July 2000
A: Reference number: Z25286
A: Accession: T51071
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-728 <SCH>
A: Cross-references: EMBL:AL390092; GSPDB: GN00116; NCSP: B2A19.50
A: Experimental source: BAC clone B2A19; strain OR74A
C: Genetics:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:08:51 ; Search time 7.57895 Seconds
(without alignments)
82.444 Million cell updates/sec

Title: US-09-444-281-36
Perfect score: 86
Sequence: 1 ILRPPWPPRRK 12

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70	81.4	144	INDC_BOVIN	P33046 bos taurus
2	53	61.6	1173	VGL2_CVR22	P15423 human coron
3	49	57.0	226	WS18_HUMAN	Q96119 homo sapien
4	49	57.0	492	ADRO_BOVIN	P08165 bos taurus
5	47	54.7	253	Y945_MYCTU	P71564 mycobacteri
6	47	54.7	715	Y945_MYCTU	Q11025 mycobacteri
7	45.5	52.9	505	TRPE_PSESS	P21689 pseudomonas
8	45	52.3	196	YA05_SCHPO	Q09677 schizosacch
9	45	52.3	1108	GN3B_RAT	Q53085 rattus norv
10	44	51.2	361	FUT3_HUMAN	P21217 homo sapien
11	44	51.2	372	FUT3_PANTR	Q19058 pan troglod
12	44	51.2	397	MM16_MYCTU	Q10773 mycobacteri
13	44	51.2	535	YD06_SCHPO	Q13912 schizosacch
14	44	51.2	967	MM14_MYCTU	Q53735 mycobacteri
15	44	51.2	968	MM12_MYCTU	Q11171 mycobacteri
16	44	51.2	985	SK13_MOUSE	Q44891 mus musculu
17	44	51.2	1154	VGL2_IBVD2	P12722 avian infec
18	44	51.2	1162	VGL2_IBVK	P11223 avian infec
19	44	51.2	1162	VGL2_IBVK	P12650 avian infec
20	44	51.2	1162	VGL2_IBVM	P12651 avian infec
21	44	51.2	1163	VGL2_IBVM	P05135 avian infec
22	44	51.2	1255	VGL2_CVRSA	P59594 human coron
23	43.5	50.6	276	RCEL_RHOFA	Q33005 rhodopsu
24	43.5	50.6	2436	ABE2_HUMAN	Q95267 homo sapien
25	43	50.0	51	LHB2_ECTHA	P80105 ectothiorho
26	43	50.0	250	NPD1_PSEAE	Q91410 pseudomonas
27	43	50.0	250	NPD1_PSEAE	Q87490 methanopyru
28	43	50.0	256	NPD4_PSESM	Q87490 pseudomonas
29	43	50.0	711	NMLA_STRCO	Q53902 streptomyc
30	43	50.0	958	NML1_MYCTU	P95211 mycobacteri
31	43	50.0	1112	GN3B_HUMAN	Q13370 homo sapien
32	43	50.0	1225	VGL2_CVPR8	P27655 porcine res
33	43	50.0	1225	VGL2_CVPRM	P24413 porcine res

ALIGNMENTS

RESULT 1

INDC_BOVIN	STANDARD;	PRT;	144 AA.
AC P33046;			
DT 01-OCT-1993 (Rel. 27, Created)			
DT 01-OCT-1993 (Rel. 27, Last sequence update)			
DT 10-OCT-2003 (Rel. 42, Last annotation update)			
DE Indolicidin precursor.			
OS Bos taurus (Bovine).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC Bovidae; Bovinae; Bos.			
OX NCBI_TaxID=9913;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC TISSUE=Bone marrow;			
RX MEDLINE=92392368; PubMed=1520337;			
RA del Sal G., Storici P., Schneider C., Romeo D., Zanetti M.;			
RT "cDNA cloning of the neutrophil bactericidal peptide indolicidin.";			
RL Biochem. Biophys. Res. Commun. 187:467-472(1992).			
RN [2]			
RP SEQUENCE OF 131-143.			
RC TISSUE=Neutrophils;			
RX MEDLINE=92165771; PubMed=1537821;			
RA Seilsted M.B., Novotny M.J., Morris W.L., Tang Y.-Q., Smith W.,			
RA Cullor J.S.;			
RT "Indolicidin, a novel bactericidal tridecapeptide amide from			
RT neutrophils.";			
RL J. Biol. Chem. 267:4292-4295(1992).			
CC -!- FUNCTION: Potent microbicidal activity, active against			
CC Staphylococcus aureus and Escherichia coli.			
CC -!- TISSUE SPECIFICITY: Large granules of neutrophils.			
CC -!- PTM: Elastase might be responsible for its maturation.			
CC -!- SIMILARITY: Belongs to the cathelicidin family.			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC the European Bioinformatics Institute. There are no restrictions on its			
CC use by non-profit institutions as long as its content is in no way			
CC modified and this statement is not removed. Usage by and for commercial			
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC or send an email to license@sib-sib.ch).			
CC -----			
CC EMBL; X67340; CAA47755.1; -			
DR FIR; JCI222; JCI222.			
DR PDB; 1G89; 17-JAN-01.			
DR PDB; 1G8C; 17-JAN-01.			
DR PDB; 1HR1; 31-DEC-02.			
DR InterPro; IPR001894; Cathelicidin.			
DR Pfam; PF00666; Cathelicidin; 1.			
DR ProDom; PD001838; Cathelicidin; 1.			
DR PROSITE; PS00946; CATHELICIDINS_1; 1.			
DR PROSITE; PS00947; CATHELICIDINS_2; 1.			
KW Antibiotic; Amidation; Signal; Pyrrolidone carboxylic acid;			
KW 3D-structure.			
FT SIGNAL	1	29	POTENTIAL.

P11225 murine coro
P11224 murine coro
P36334 human coro
P25190 bovine coro
P25191 bovine coro
P25192 bovine coro
P15777 bovine coro
P25193 bovine coro
P25194 bovine coro
P22432 murine coro
Q02385 murine coro
P07946 porcine tra


```

FT PROPEP 30 130
FT PEPTIDE 131 143
FT MOD_RES 30 30
FT ID VGL2_CVH22 STANDARD; PRT: 1173 AA.
AC P15423; P89342; P89343; P89344; Q86174; Q990M1; Q990M2; Q990M3;
AC Q990M4;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein).
GN S.
OS Human coronavirus (strain 229E) (HCoV-229E).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11137;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=90264837; PubMed=2345357;
RX Raabe T., Schelle-Prinz B., Siddell S.G.;
RA "Nucleotide sequence of the gene encoding the spike glycoprotein of
RT human coronavirus HCV 229E."
RL J. Gen. Virol. 71:1065-1073(1990).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=21262210; PubMed=11369870;
RA Thiel V., Herold J., Schelle B., Siddell S.G.;
RT "Infectious RNA transcribed in vitro from a cDNA copy of the human
RL coronavirus genome cloned in vaccinia virus."
RN J. Gen. Virol. 82:1273-1281(2001).
RN [3]
RN SEQUENCE OF 98-1113 FROM N.A., AND VARIANTS.
RC STRAIN=Isolate RW Stock, Isolate F100E, Isolate P1A, and
RC Isolate P1B;
RA Bonavia A., Holmes K.V.;
RT "Viral and cellular changes in a human cell line persistently infected
RL with human coronavirus HCoV-229E."
RN Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE OF 98-1113 FROM N.A., AND VARIANTS.
RC STRAIN=Isolate ATCC VR-74, Isolate A162, and Isolate LRI 281;
RX MEDLINE=99086140; PubMed=9870593;
RA Hays J.P., Myint S.H.;
RT "PCR sequencing of the spike genes of geographically and
RL chronologically distinct human coronaviruses 229E."
RN J. Virol. Methods 75:179-193(1998).
RN [5]
RN SEQUENCE OF 1159-1173 FROM N.A.
RX MEDLINE=89366667; PubMed=2701946;
RA Raabe T., Siddell S.;
RT "Nucleotide sequence of the human coronavirus HCV 229E mRNA 4 and mRNA
RL 5 unique regions."
RN Nucleic Acids Res. 17:6387-6387(1989).
RN [6]
RN INTERACTION WITH ANPEP.
RX MEDLINE=22440020; PubMed=12551991;
RA Bonavia A., Zelus B.D., Wentworth D.E., Talbot P.J., Holmes K.V.;

```

```

RT "Identification of a receptor-binding domain of the spike glycoprotein
of human coronavirus HCoV-229E."
J. Virol. 77:2530-2538(2003).
RN [7]
RN INTERACTION WITH ANPEP.
RX MEDLINE=22521439; PubMed=12634402;
RA Breslin J.J., Mork I., Smith M.K., Vogel L.K., Hemmila E.M.,
RA Bonavia A., Talbot P.J., Sjoestrom H., Noren O., Holmes K.V.;
RT "Human coronavirus 229E: receptor binding domain and neutralization by
RL soluble receptor at 37 degrees C."
RN J. Virol. 77:4435-4438(2003).
RN [8]
RN REVIEW
RX MEDLINE=21109095; PubMed=11162792;
RA Gallagher T.M., Buchmeier M.J.;
RT "Coronavirus spike proteins in viral entry and pathogenesis."
RL Virology 279:371-374(2001).
CC -|- FUNCTION: Structural protein that makes spikes at the surface of
the virus. Determines enteropathogenicity and virulence of the
virus. Initiates infection by specifically recognizing and binding
the human aminopeptidase ANPEP receptor. Its association with
ANPEP may lead to its conformational change that triggers fusion
between viral and host cellular membrane.
CC -|- SUBUNIT: Homotrimer. During virus morphogenesis, it is found in a
complex with M and HE proteins (By similarity). Interacts with
ANPEP.
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
CC -|- DOMAIN: The spike S1 domain displays the specificity for the host
receptor.
CC -|- DOMAIN: The leucine zipper-like heptad repeats may mediate the
fusion of viral and cellular membranes.
CC -|- POLYMORPHISM: The strong variation between the different
strains may affect the virulence of the virus.
CC -|- MISCELLANEOUS: In contrast to serogroup 2, E2 glycoprotein protein
from serogroup 1 is not cleaved.
CC -|- SIMILARITY: Contains 1 spike S1 domain.
CC -|- SIMILARITY: Contains 1 spike S2 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
CC -----
EMBL; X16816; CAA34723.1; -
EMBL; AF304450; AAG48592.1; -
EMBL; AF344186; AAK32188.1; -
EMBL; AF344187; AAK32189.1; -
EMBL; AF344188; AAK32190.1; -
EMBL; AF344189; AAK32191.1; -
EMBL; Y09923; CAA71056.1; -
EMBL; Y10051; CAA71146.1; -
EMBL; Y10052; CAA71147.1; -
EMBL; X15654; CAA33680.1; -
PIR; A34766; VGIHHC.
DR InterPro; IPR002551; Corona_S1.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01600; Corona_S1; 1.
DR Pfam; PF01601; Corona_S2; 1.
DR Virulence; Glycoprotein; Envelope protein; Transmembrane; Signal;
KW Coiled coil.
FT SIGNAL 1 15
FT CHAIN 16 1173
FT DOMAIN 16 1115
FT TRANSMEM 1116 1173
FT DOMAIN 1136 1173
FT DOMAIN 32 536
FT DOMAIN 417 547
FT DOMAIN 537 1171
FT DOMAIN 1054 1103
FT DOMAIN 1067 1102

```

FT	DOMAIN	1136	1157	CYS-RICH.	FT	VARIANT	775	A -> S (in isolate A162).
FT	CARBOHYD	23	23	N-LINKED (GLCNAC. . .)	FT	WS18 HUMAN	775	61.6%; Score 53; DB 1; Length 1173;
FT	CARBOHYD	62	62	N-LINKED (GLCNAC. . .)	FT	AC Q96LE9; Q9BSG8;		Best Local Similarity 62.5%; Pred. No. 7.5;
FT	CARBOHYD	98	98	N-LINKED (GLCNAC. . .)	FT	DT 28-FEB-2003 (Rel. 41, Created)		Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0
FT	CARBOHYD	147	147	N-LINKED (GLCNAC. . .)	FT	DT 28-FEB-2003 (Rel. 41, Last sequence update)		
FT	CARBOHYD	171	171	N-LINKED (GLCNAC. . .)	FT	DE 10-OCT-2003 (Rel. 42, Last annotation update)		
FT	CARBOHYD	176	176	N-LINKED (GLCNAC. . .)	FT	DE Williams-Beuren syndrome chromosome region 18 protein.		
FT	CARBOHYD	220	220	N-LINKED (GLCNAC. . .)	FT	GN WBSR18.		
FT	CARBOHYD	243	243	N-LINKED (GLCNAC. . .)	FT	OS Homo sapiens (Human)		
FT	CARBOHYD	326	326	N-LINKED (GLCNAC. . .)	FT	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
FT	CARBOHYD	333	333	N-LINKED (GLCNAC. . .)	FT	OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
FT	CARBOHYD	440	440	N-LINKED (GLCNAC. . .)	FT	OX NCBI_Taxid=9606;		
FT	CARBOHYD	464	464	N-LINKED (GLCNAC. . .)	FT	RN [1]		
FT	CARBOHYD	518	518	N-LINKED (GLCNAC. . .)	FT	RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.		
FT	CARBOHYD	538	538	N-LINKED (GLCNAC. . .)	FT	RX MEDLINE=22067697; PubMed=12073013;		
FT	CARBOHYD	542	542	N-LINKED (GLCNAC. . .)	FT	RA Merla G., Ucla C., Guipponi M., Raymond A.;		
FT	CARBOHYD	568	568	N-LINKED (GLCNAC. . .)	FT	RT "Identification of additional transcripts in the Williams-Beuren		
FT	CARBOHYD	581	581	N-LINKED (GLCNAC. . .)	FT	RL Hum. Genet. 110:429-438(2002).		
FT	CARBOHYD	663	663	N-LINKED (GLCNAC. . .)	FT	RN [2]		
FT	CARBOHYD	671	671	N-LINKED (GLCNAC. . .)	FT	RP SEQUENCE FROM N.A.		
FT	CARBOHYD	930	930	N-LINKED (GLCNAC. . .)	FT	RC TISSUE=Testis;		
FT	CARBOHYD	1015	1015	N-LINKED (GLCNAC. . .)	FT	RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,		
FT	CARBOHYD	1020	1020	N-LINKED (GLCNAC. . .)	FT	RA Houta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J., Chiba Y.,		
FT	CARBOHYD	1037	1037	N-LINKED (GLCNAC. . .)	FT	RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,		
FT	CARBOHYD	1049	1049	N-LINKED (GLCNAC. . .)	FT	RA Suzuki Y., Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,		
FT	CARBOHYD	1061	1061	N-LINKED (GLCNAC. . .)	FT	RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,		
FT	CARBOHYD	1066	1066	N-LINKED (GLCNAC. . .)	FT	RA Kawakami B., Nagai K., Isogai T., Sugano S.;		
FT	CARBOHYD	1076	1076	N-LINKED (GLCNAC. . .)	FT	RT "NEDO human cDNA sequencing project."		
FT	CARBOHYD	1082	1082	N-LINKED (GLCNAC. . .)	FT	RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.		
FT	CARBOHYD	1096	1096	N-LINKED (GLCNAC. . .)	FT	RN [3]		
FT	VARIANT	98	98	N -> S (in isolate LRI 281).	FT	RC TISSUE=Lung;		
FT	VARIANT	120	120	N -> I (in isolate LRI 281).	FT	RX MEDLINE=22388257; PubMed=12477932;		
FT	VARIANT	127	127	LR -> IS (in isolate A162).	FT	RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
FT	VARIANT	176	176	N -> T (in isolate P100E).	FT	RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
FT	VARIANT	210	210	T -> S (in isolate A162).	FT	RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
FT	VARIANT	223	223	T -> N (in isolate A162).	FT	RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,		
FT	VARIANT	228	228	DF -> V (in isolate A162).	FT	RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
FT	VARIANT	230	230	C -> F (in isolate LRI 281).	FT	RA Staphenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
FT	VARIANT	230	230	C -> L (in isolates RW Stock, P11A, P11B, P100E and ATCC VR-74).	FT	RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
FT	VARIANT	248	248	S -> A (in isolate A162).	FT	RA Raha S.S., Loquellanc N.A., Peters G.J., Abramson R.D., Mullany S.J.,		
FT	VARIANT	270	270	D -> Y (in isolate P100E).	FT	RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
FT	VARIANT	295	295	V -> A (in isolate A162).	FT	RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
FT	VARIANT	300	300	T -> M (in isolate P100E).	FT	RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
FT	VARIANT	307	307	D -> N (in isolate A162).	FT	RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,		
FT	VARIANT	310	311	PO -> LR (in isolate A162).	FT	RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
FT	VARIANT	314	324	GKGCNCYCPAG -> VGRYCNCRAY (in isolate A162).	FT	RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
FT	VARIANT	336	336	K -> N (in isolate LRI 281).	FT	RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,		
FT	VARIANT	349	358	KTVAVYANGV -> QFVGAKRD (in isolate A162).	FT	RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.B.,		
FT	VARIANT	401	401	V -> N (in isolate A162).	FT	RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.;		
FT	VARIANT	404	411	WAYSKYVT -> LAMLNSHN (in isolate A162).	FT	RT "Generation and initial analysis of more than 15,000 full-length		
FT	VARIANT	414	414	S -> T (in isolate P100E).	FT	RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
FT	VARIANT	424	424	G -> V (in isolate A162).	FT	CC -!- TISSUE SPECIFICITY: Expressed in brain, heart, kidney, liver,		
FT	VARIANT	430	430	Q -> K (in isolate A162).	FT	CC lung, spleen, stomach and testis.		
FT	VARIANT	441	441	V -> A (in isolate LRI 281).	FT	CC -!- DISEASE: Haploinsufficiency of WBSR18 may be the cause of certain		
FT	VARIANT	444	444	D -> N (in isolate A162).	FT	CC cardiovascular and musculo-skeletal abnormalities observed in		
FT	VARIANT	462	462	V -> I (in isolate A162).	FT	CC Williams-Beuren syndrome (WBS), a rare developmental disorder. It		
FT	VARIANT	481	481	L -> N (in isolate A162).	FT	CC is a contiguous gene deletion syndrome involving genes from		
FT	VARIANT	488	488	K -> V (in isolate A162).	FT	CC chromosome band 7q11.23.		
FT	VARIANT	530	530	L -> M (in isolate A162).	FT			
FT	VARIANT	577	577	I -> T (in isolate P11A).	FT			
FT	VARIANT	578	578	T -> G (in isolate P100E).	FT			
FT	VARIANT	590	590	R -> M (in isolate A162).	FT			
FT	VARIANT	642	642	R -> I (in isolate A162).	FT			
FT	VARIANT	681	681	L -> R (in isolate A162).	FT			
FT	VARIANT	700	700	L -> I (in isolates RW Stock, P11A, P11B and P100E).	FT			
FT	VARIANT	711	711	K -> N (in isolate LRI 281).	FT			
FT	VARIANT	714	714	K -> N (in isolates RW Stock, P11A, P11B and P100E).	FT			
FT	VARIANT	765	765	V -> A (in isolate A162).	FT			

```

CC  -!- SIMILARITY: Contains 1 J domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF412025; AAM62307.1; -.
DR  EMBL; AK058113; BAB71671.1; -.
DR  EMBL; BC005056; AAK05056.1; -.
DR  Genew; HGNC:16410; WBSCR18.
DR  InterPro; IPR001623; DnaJ_N.
DR  Pfam; PF00226; DnaJ; 1.
DR  PRINTS; P00625; DnaJPROTEIN.
DR  SMART; SM00271; DnaJ; 1.
DR  PROSITE; PS00636; DnaJ_1; FALSE_NEG.
DR  PROSITE; PS0076; DnaJ_2; 1.
KW  Chaperone; Williams-Beuren syndrome.
FT  DOMAIN 49 114 J-DOMAIN.
FT  CONFLICT 34 34 G -> R (IN REF. 3).
SQ  SEQUENCE 226 AA; 25961 MW; 8687C2A45790381D CRC64;

Query Match 57.0%; Score 49; DB 1; Length 226;
Best Local Similarity 53.8%; Pred. No. 5.2;
Matches 7; Conservative 1; Mismatches 1; Indels 4; Gaps 1;

QY  2 LRWPMW----PWR 10
DB  :||| ||| |||
    4 MNRWQRLLPWR 16

RESULT 4
ADRO BOVIN STANDARD; PRT; 492 AA.
AC  P08165; Q95KN8.
DT  01-AUG-1988 (Rel. 08, Created)
DT  15-JUL-1998 (Rel. 36, Last sequence update)
DT  15-MAR-2004 (Rel. 43, Last annotation update)
DE  NADPH:adrenodoxin oxidoreductase, mitochondrial precursor
DE  (SC 1.18.1.2) (Adrenodoxin reductase) (AR) (Ferredoxin-NADP(-)
DE  reductase).
DE  FDXR OR ADXR.
GN  Bos taurus (Bovine).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC  Bovidae; Bovinae; Bos.
OX  NCBI_TaxID=9913;
RN  [1]
RP  SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX  MEDLINE=94177140; PubMed=8130767;
RA  Takata Y., Sagara Y., Kono A., Sekimizu K., Horiuchi T.;
RT  "Gene structure of bovine adrenodoxin reductase.";
RL  Biol. Pharm. Bull. 16:1200-1206(1993).
RN  [2]
RP  SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX  MEDLINE=88198050; PubMed=3448086;
RA  Sagara Y., Takata Y., Miyata T., Hara T., Horiuchi T.;
RT  "Cloning and sequence analysis of adrenodoxin reductase cDNA from
RT  bovine adrenal cortex.";
RL  J. Biochem. 102:1333-1336(1987).
RN  [3]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=87270696; PubMed=3038094;
RA  Nonaka Y., Murakami H., Yabusaki Y., Kuramitsu S., Kagamiyama H.,
RA  Yamano T., Okamoto M.;
RT  "Molecular cloning and sequence analysis of full-length cDNA for mRNA
RT  of adrenodoxin oxidoreductase from bovine adrenal cortex.";
RL  Biochem. Biophys. Res. Commun. 145:1239-1247(1987).
RN  [4]

```

```

RP  SEQUENCE FROM N.A.
RC  TISSUE=Adrenal cortex;
RX  MEDLINE=89170752; PubMed=2924777;
RA  Hanukoglu I., Gutfinger T.;
RT  "cDNA sequence of adrenodoxin reductase. Identification of NADP-
RT  binding sites in oxidoreductases.";
RL  Eur. J. Biochem. 180:479-484(1989).
RN  [5]
RP  SEQUENCE OF N-TERMINUS, AND PARTIAL SEQUENCE.
RC  TISSUE=Adrenal cortex;
RX  MEDLINE=88082777; PubMed=3691502;
RA  Hanukoglu I., Gutfinger T., Hanu M., Shively J.E.;
RT  "Isolation of a cDNA for adrenodoxin reductase (ferredoxin-NADP+
RT  reductase). Implications for mitochondrial cytochrome P-450 systems.";
RL  Eur. J. Biochem. 169:449-455(1987).
RN  [6]
RP  X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 33-492.
RC  TISSUE=Adrenal gland;
RX  MEDLINE=99299392; PubMed=10369776;
RA  Ziegler G.A., Vornheim C., Hanukoglu I., Schulz G.E.;
RT  "The structure of adrenodoxin reductase of mitochondrial P450 systems:
RT  electron transfer for steroid biosynthesis.";
RL  J. Mol. Biol. 289:981-990(1999).
RN  [7]
RP  X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
RX  MEDLINE=20455764; PubMed=10998235;
RA  Ziegler G.A., Schulz G.E.;
RT  "Crystal structures of adrenodoxin reductase in complex with NADP+ and
RT  NADPH suggesting a mechanism for the electron transfer of an enzyme
RT  family.";
RL  Biochemistry 39:10986-10995(2000).
RN  [8]
RP  X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH ADRENODOXIN.
RX  MEDLINE=21264735; PubMed=11053423;
RA  Mueller J.J., Lapko A., Bourenkov G., Ruckpaul K., Heinemann U.;
RT  "Adrenodoxin reductase-adrenodoxin complex structure suggests electron
RT  transfer path in steroid biosynthesis.";
RL  J. Biol. Chem. 276:2786-2789(2001).
CC  -!- FUNCTION: Serves as the first electron transfer protein in all the
CC  mitochondrial P450 systems. Including cholesterol side chain
CC  cleavage in all steroidogenic tissues, steroid 11-beta
CC  hydroxylation in the adrenal cortex, 25-OH-vitamin D3-24
CC  hydroxylation in the kidney, and sterol C-27 hydroxylation in the
CC  liver.
CC  -!- CATALYTIC ACTIVITY: Reduced adrenodoxin + NADP(+) = oxidized
CC  adrenodoxin + NADPH.
CC  -!- COFACTOR: FAD.
CC  -!- PATHWAY: Cholesterol side-chain-cleavage system.
CC  -!- SUBUNIT: Monomer.
CC  -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC  -!- ALTERNATIVE PRODUCTS:
CC  Event=Alternative splicing; Named isoforms=2;
CC  Name=Short;
CC  IsoId=P08165-1; Sequence=Displayed;
CC  Name=Long;
CC  IsoId=P08165-2; Sequence=VSP_003415;
CC  Note=Represents 10-20% of all adrenodoxin reductase mRNAs and
CC  seems to be inactive;
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; D83475; BAA11921.1; -.
DR  EMBL; D83472; BAA11921.1; JOINED.
DR  EMBL; D83473; BAA11921.1; JOINED.
DR  EMBL; D83474; BAA11921.1; JOINED.
DR  EMBL; M17029; AAA00362.1; -.
DR  EMBL; D00211; AAA00150.1; -.

```

```

DR EMBL; X13736; CAA32002.1; -.
DR PIR; JT0751; JT0751.
DR PDB; 1CJC; 12-APR-99.
DR PDB; 1E1L; 24-SEP-00.
DR PDB; 1E1K; 24-SEP-00.
DR PDB; 1E1M; 24-SEP-00.
DR PDB; 1E1N; 24-SEP-00.
DR PDB; 1E6E; 01-AUG-03.
DR InterPro; IPR000759; Adrndx_reductase.
DR PRINTS; PR00419; ADXRDTASE.
KW Electron transport; Cholesterol metabolism; Oxidoreductase;
KW Mitochondrion; FAD; Flavoprotein; NADP; Transit peptide;
KW Alternative splicing; 3D-structure.
FT TRANSIT 1 32 MITOCHONDRION.
FT CHAIN 33 492 NADPH:ADRENODOXIN OXIDOREDUCTASE.
FT VARSPLIC 204 204 E -> EVLLLCQ (in isoform Long).
FT /FTid=VSP_003415.
FT CONFLICT 77 77 G -> R (IN REF. 3).
FT CONFLICT 81 94 FGVAPDHPVKNVI -> VMLALTPSRMLL (IN REF. 3).
FT CONFLICT 124 128 QDAVH -> RVYRLT (IN REF. 3).
FT CONFLICT 268 268 K -> R (IN REF. 3).
FT CONFLICT 317 318 PS -> RL (IN REF. 3).
FT CONFLICT 323 333 RAAGIRLAVTR -> ARRSAQSP (IN REF. 3).
FT CONFLICT 341 352 TRAVPTGDVEDL -> HPGSAHWGCGGP (IN REF. 3).
FT STRAND 40 44
FT HELIX 48 60
FT STRAND 65 69
FT TURN 77 77
FT HELIX 78 81
FT TURN 82 82
FT TURN 85 86
FT HELIX 88 92
FT TURN 102 102
FT TURN 104 105
FT STRAND 110 110
FT STRAND 114 114
FT TURN 115 117
FT STRAND 118 118
FT HELIX 120 126
FT STRAND 129 132
FT STRAND 138 139
FT TURN 145 148
FT TURN 150 151
FT STRAND 152 154
FT HELIX 155 162
FT TURN 163 164
FT TURN 166 168
FT TURN 169 170
FT TURN 175 176
FT STRAND 179 183
FT HELIX 187 197
FT HELIX 200 203
FT TURN 204 205
FT HELIX 210 217
FT TURN 218 218
FT STRAND 223 227
FT HELIX 232 234
FT TURN 239 246
FT TURN 247 247
FT TURN 249 250
FT STRAND 251 254
FT HELIX 257 260
FT TURN 261 262
FT HELIX 263 266
FT TURN 267 269
FT HELIX 272 286
FT TURN 291 298
FT TURN 299 299
FT STRAND 302 307
FT STRAND 310 317
FT TURN 319 320

Query Match 57.0%; Score 49; DB 1; Length 492;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WPWEPW 9
Db 6 WRWEPW 11

RESULT 5
Y945_MYCTU STANDARD; PRT; 253 AA.
AC P71564;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative oxidoreductase K0945/M10971 (EC 1.-.-.-).
GN RV0945 OR M10971 OR MTCY10D7.29C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Broech R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekai F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Knouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.W.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC

```

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; Z79700; CAB02005.1; --
 CC EMBL; AB06982; AAK45215.1; --
 CC PIR; G70715; G70715.
 CC TIGR; MT0971; --
 CC TubercuList; RV0945; --
 CC InterPro; IPR002198; ADH short.
 CC Pfam; PF00106; adh short; 1.
 CC PRINTS; PG00080; SDRFAMILY.
 CC PROSITE; PS00061; ADH_SHORT; 1.
 CC Hypothetical protein; Oxidoreductase; Complete proteome.
 KW Hypothetical protein; Oxidoreductase; Complete proteome.
 FT ACT SITE 159 159 BY SIMILARITY.
 SQ SEQUENCE 253 AA; 27138 MW; BAD937208942DA12 CRC64;

Query Match 54.7%; Score 47; DB 1; Length 253;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PWWPW 9
 |||||
 Db 230 PWWPW 234

RESULT 6

ID YD55 MYCTU STANDARD; PRT; 715 AA.
 AC Q11025;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein RV1355C/MT1398.
 GN RV1355C OR MT1398 OR MTC102B10.19C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634220;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaiia F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence";
 RL Nature 393:537-544 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=2206494; PubMed=12218036;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J.C., Khouri H., Gill J., Mikula A.,
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
 RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains";
 RL J. Bacteriol. 184:5479-5490 (2002).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; Z75555; CAA39988.1; --
 CC EMBL; AB07012; AAK45661.1; ALT_INIT.
 CC PIR; B70741; B70741.
 CC TIGR; MT1398; --
 CC TubercuList; RV1355C; --
 CC InterPro; IPR003036; Moeb.
 CC InterPro; IPR000594; Th1F_domain.
 CC Pfam; PF00899; Th1F; 1.
 CC KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 715 AA; 78181 MW; 455495248A56041C CRC64;

Query Match 54.7%; Score 47; DB 1; Length 715;
 Best Local Similarity 66.7%; Pred. No. 28;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWPWPWR 11
 |||||
 Db 65 RWAYPWR 73

RESULT 7

ID TRPE PSSE STANDARD; PRT; 505 AA.
 AC P21689;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Anthranilate synthase component I (EC 4.1.3.27).
 GN TRPE.
 OS Pseudomonas syringae (pv. savastanoi).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=29438;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91100331; PubMed=1987141;
 RA da Costa e Silva O., Kosuge T.;
 RA "Molecular characterization and expression analysis of the
 RT anthranilate synthase gene of Pseudomonas syringae subsp.
 RT savastanoi";
 RL J. Bacteriol. 173:463-471 (1991).
 CC -1- CARBAMYLIC ACTIVITY: Chorismate + L-glutamine = anthranilate +
 CC pyruvate + L-glutamate.
 CC -1- PATHWAY: Tryptophan biosynthesis; first step.
 CC -1- SUBUNIT: Tetramer of two components I and two components II (By
 CC similarity).
 CC -1- MISCELLANEOUS: Component I catalyzes the formation of anthranilate
 CC using ammonia rather than glutamine, whereas component II provides
 CC glutamine amidotransferase activity.
 CC -1- SIMILARITY: Belongs to the anthranilate synthase component I
 CC family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M55911; AAA26016.1; --
 CC HSP; Q06128; 1QDL
 CC InterPro; IPR005801; Anth_synth_chor.
 CC InterPro; IPR006805; Anth_synth_N.
 CC InterPro; IPR005256; Anth_synthI.
 CC Pfam; PF04715; Anth_synth_I_N; 1.
 CC Pfam; PF00425; chorismate bind; 1.
 CC PRINTS; PR00095; ANTSNTHASEI.
 CC ProDom; PD000779; Anth_synth_chor; 1.

KW Hypothetical protein.
SQ SEQUENCE 196 AA; 22104 MW; 436764DA9E26074C CRC64;
Query Match 52.3%; Score 45; DB 1; Length 196;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 8; Conservative 2; Mismatches 2; Indels 4; Gaps 2;
QY 1 ILRWP-KW---PWRK 12
DB 63 IYQPGWMTGTPWKLK 78
RESULT 9
CN3B_RAT
ID CN3B_RAT STANDARD; PRT; 1108 AA.
AC Q63085;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE GMP-inhibited 3',5'-cyclic phosphodiesterase B (EC 3.1.4.17) (Cyclic
DE GMP inhibited phosphodiesterase B) (CGI-PDE B) (CGIPDE1).
GN PDE3B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE=Adipose tissue;
RX MEDLINE=93366761; PubMed=895509;
RA Taira M., Hockman S.C., Calvo J.C., Taira M., Belfrage P.,
RA Manganiello V.C.;
RT Inhibited cyclic nucleotide phosphodiesterase.";
RL J. Biol. Chem. 268:18573-18579(1993).
CC -!- FUNCTION: May play a role in fat metabolism.
CC -!- CATALYTIC ACTIVITY: Nucleoside 3',5'-cyclic phosphate + H(2)O =
CC nucleoside 5'-phosphate.
CC -!- ENZYME REGULATION: Inhibited by cGMP.
CC -!- SUBCELLULAR LOCATION: Membrane-bound (Potential).
CC -!- TISSUE SPECIFICITY: Abundant in adipose tissues.
CC -!- SIMILARITY: Belongs to the cyclic nucleotide phosphodiesterase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 222867; CAA80489.1; -.
CC PIR; A48508; A48508.
CC InterPro; IPR003607; Met_phosphohydro.
CC InterPro; IPR003073; PDEase.
CC Pfam; PF00233; PDEase; 1.
CC SMART; SM00471; HGC; 1.
CC PROSITE; PS00126; PDEASE_1; 1.
CC HYDROLASE; cGMP; Membrane.
KW DOMAIN 16 22 POLY-PRO.
FT DOMAIN 99 102 POLY-ALA.
FT DOMAIN 175 179 POLY-ALA.
FT DOMAIN 1007 1021 POLY-ASP.
FT DOMAIN 1068 1071 POLY-GLU.
FT DOMAIN 1101 1104 POLY-GLU.
SQ SEQUENCE 1108 AA; 123105 MW; C9B5078C7D3ADD6D CRC64;
Query Match 52.3%; Score 45; DB 1; Length 1108;
Best Local Similarity 62.5%; Pred. No. 76;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 4 WFWPWR 11

DR TIGRfams; TIGR00564; trpE_most; 1.
SQ SEQUENCE 505 AA; 56084 MW; A38E819331331F6BB CRC64;
Query Match 52.9%; Score 45.5; DB 1; Length 505;
Best Local Similarity 28.0%; Pred. No. 31;
Matches 7; Conservative 2; Mismatches 3; Indels 13; Gaps 1;
QY 1 ILRW-----PWWPWRK 12
DB 467 VLEWEETLNKRRAMVGSAAWPPRR 491
RESULT 8
YAO5_SCHPO
ID YAO5_SCHPO STANDARD; PRT; 196 AA.
AC Q09677;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein CSH10.05C in chromosome I.
GN SPACSH10.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Guilliam R., Rajandream M.A., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Murgall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard J.C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Foreburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- SIMILARITY: STRONG, TO BACTERIAL MODULATOR OF DRUG ACTIVITY B
CC (MDAB).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 249811; CAA89955.1; -.
CC PIR; S55483; S55483.
CC GeneDB Spombe; SPACSH10.05C; -.
CC InterPro; IPR003680; NADPHdh 2.
CC Pfam; PF02525; Flavodoxin 2.
CC

Db 164 WQWNSLR 171

RESULT 10

ID FUT3 HUMAN STANDARD; PRT; 361 AA.

AC P21217; Q99448; Q99449;

DT 01-MAY-1991 (Rel. 18, Created)

DT 01-MAY-1991 (Rel. 18, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Galactoside 3(4)-L-fucosyltransferase (EC 2.4.1.65) (Blood group Lewis

DE alpha-4-fucosyltransferase) (Lewis Fx) (Fucosyltransferase 3) (FUCT-

DE III) OR LE OR FUT3B.

GN FUT3 OR LE OR FUT3B.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1] _SEQUENCE FROM N.A.

RP MEDLINE=91032981; PubMed=1977660;

RX Kowaska-Latallo J.F., Larsen R.D., Nair R.P., Lowe J.B.;

RA "A cloned human cDNA determines expression of a mouse stage-specific

RT embryonic antigen and the Lewis blood group

RT alpha(1,3/1,4)fucosyltransferase.";

RL Genes Dev. 4:1288-1303(1990).

RN [2] _SEQUENCE FROM N.A.

RP MEDLINE=95378269; PubMed=7650030;

RX Tissue-Liver;

RA Cameron H.S., Szczepaniak D., Weston W.;

RT "Expression of human chromosome 19p alpha(1,3)-fucosyltransferase

RT genes in normal tissues. Alternative splicing, polyadenylation, and

RT isoforms.";

RL J. Biol. Chem. 270:20112-20122(1995).

RN [3] _SEQUENCE FROM N.A.

RP MEDLINE=94059067; PubMed=8240322;

RX Elmgren A., Rydberg L., Larson G.;

RA "Genotypic heterogeneity among Lewis negative individuals.";

RT Biochem. Biophys. Res. Commun. 196:515-520(1993).

RN [5] _SEQUENCE FROM N.A.

RP MEDLINE=94059082; PubMed=8240337;

RX Nishihara S., Yazawa S., Iwasaki H., Nakazato M., Kudo T., Ando T.,

RA Narimatsu H.;

RT "Alpha (1,3/1,4)fucosyltransferase (FucT-III) gene is inactivated by

RT a single amino acid substitution in Lewis histo-blood type negative

RT individuals.";

RL Biochem. Biophys. Res. Commun. 196:624-631(1993).

RN [6] _SEQUENCE FROM N.A.

RP MEDLINE=94033579; PubMed=8219240;

RX Koda Y., Kimura H., Nekada E.;

RA "Analysis of Lewis fucosyltransferase genes from the human gastric

RT mucosa of Lewis-positive and -negative individuals.";

RL Blood 82:2915-2919(1993).

RN [7] _SEQUENCE FROM N.A.

RP MEDLINE=94342599; PubMed=8063716;

RX Mollicone R., Reuguine I., Kelly R.J., Fletcher A., Watt J.,

RA Chatfield S., Aziz A., Cameron H.S., Weston B.W., Lowe J.B., Oriol R.;

RT "Molecular basis for Lewis alpha(1,3/1,4)-fucosyltransferase gene

RT deficiency (FUT3) found in Lewis-negative Indonesian pedigrees.";

RL J. Biol. Chem. 269:20987-20994(1994).

RN [8] _VARIANT LE(-) LYS-356.

RP MEDLINE=95050753; PubMed=7961837;

RX Nishihara S., Narimatsu H., Iwasaki H., Yazawa S., Akamatsu S.,

RA Ando T., Seno T., Narimatsu I.;

RT "Molecular genetic analysis of the human Lewis histo-blood group

RT system.";

RL J. Biol. Chem. 269:29271-29278(1994).

RN [9] _VARIANT LE(-) ARG-20; ARG-68; MET-105 AND LYS-356.

RP MEDLINE=96243526; PubMed=8801770;

RX Elmgren A., Boerjeson C., Svensson L., Rydberg L., Larson G.;

RA "DNA sequencing and screening for point mutations in the human Lewis

RT 'FUT3' gene enables molecular genotyping of the human Lewis blood

RT group system.";

RL Vox Sang. 70:97-103(1996).

RN [10] _VARIANT LE(-) ARG-68 AND MET-105.

RP MEDLINE=97413801; PubMed=9268337;

RX Elmgren A., Mollicone R., Costache M., Boerjeson C., Oriol R.,

RA Harrington J., Larson G.;

RT "Significance of individual point mutations, T202C and C314T, in the

RT human Lewis 'FUT3' gene for expression of Lewis antigens by the human

RT alpha(1,3/1,4)-fucosyltransferase, Fuc-TIII.";

RL J. Biol. Chem. 272:21994-21998(1997).

RN [11] _VARIANT LE(+) LYS-102 AND ALA-124, AND VARIANTS LE(-) ASN-162;

RP ARG-223 AND MET-270.

RX MEDLINE=98366989; PubMed=9703429;

RA Pang H., Liu Y., Koda Y., Soejima M., Jia J., Schlaphoff T.,

RA du Toit E.D., Kimura H.;

RT "Five novel missense mutations of the Lewis gene 'FUT3' in African

RT 'Xhosa' and Caucasian populations in South Africa.";

RL Hum. Genet. 102:675-680(1998).

CC -!- FUNCTION: MAY CATALYZE ALPHA-1,3 AND ALPHA-1,4 GLYCOSIDIC LINKAGES

CC INVOLVED IN THE EXPRESSION OF VIM-2, LEWIS A, LEWIS B, SIALYL

CC LEWIS X AND LEWIS X/SSER-1 ANTIGENS. MAY BE INVOLVED IN BLOOD

CC GROUP LEWIS DETERMINATION; LEWIS-POSITIVE (LE(+)) INDIVIDUALS

CC HAVE AN ACTIVE ENZYME WHILE LEWIS-NEGATIVE (LE(-)) INDIVIDUALS

CC HAVE AN INACTIVE ENZYME.

CC -!- CATALYTIC ACTIVITY: GDP-beta-L-fucose + beta-D-galactosyl-(1->3)-

CC N-acetyl-D-glucosaminyl-R = GDP + beta-D-galactosyl-(1->3)-[alpha-

CC L-fucosyl-(1->4)]-N-acetyl-beta-D-glucosaminyl-R.

CC -!- PATHWAY: Glycosylation.

CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Membrane-bound

CC form in trans cisternae of Golgi.

CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN STOMACH, COLON, SMALL

CC INTESTINE, LUNG AND KIDNEY AND TO A LESSER EXTENT IN SALIVARY

CC GLAND, BLADDER, UTERUS AND LIVER.

CC -!- MISCELLANEOUS: Also acts on the corresponding 1,4-galactosyl

CC derivative, forming 1,3-L-fucosyl links.

CC -!- SIMILARITY: Belongs to the glycosyltransferase family 10.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.

CC -----

CC EMBL; X53578; CAA37641.1; -

CC EMBL; U27328; AAC50187.1; -

CC EMBL; U27326; AAC50185.1; -

CC EMBL; U27327; AAC50186.1; -

CC EMBL; D89324; BAA13941.1; -

CC EMBL; D89325; BAA13942.1; -

CC EMBL; AF131913; AAD33514.1; -

CC PIR; A36669; A36669.

CC Genew; HGNC:4014; FUT3.

CC MIM; 111100; -

CC GO; GO:0005624; C:membrane fraction; TAS.

```

DR GO:0008417; F-fucosyltransferase activity; TAS.
DR GO:0005975; P-carbohydrate metabolism; TAS.
DR InterPro; IPR001503; Glyco_transf_10.
DR Pfam; PF00852; Glyco_transf_10; 1.
DR Transfase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack; Polymorphism; Blood group antigen.
FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 16 34 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 35 361 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (PROBABLE).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (PROBABLE).
FT VARIANT 20 20 L -> R (IN LE(-)).
FT VARIANT 68 68 W -> R (IN LE(-)).
FT VARIANT 102 102 Q -> K (IN LE(+)).
FT VARIANT 105 105 T -> M (IN LE(-)).
FT VARIANT 124 124 S -> A (IN LE(+)).
FT VARIANT 162 162 D -> N (IN LE(-)).
FT VARIANT 170 170 G -> S (IN LE(-); COMPLETELY INACTIVE).
FT VARIANT 223 223 G -> R (IN LE(-)).
FT VARIANT 270 270 V -> M (IN LE(-)).
FT VARIANT 336 336 D -> A (IN LE(-)).
FT VARIANT 356 356 I -> K (IN LE(-); LESS THAN 10% REDUCTION
IN ACTIVITY).
FT SEQUENCE 361 AA; 42117 MW; BF4398044F19C284 CRC64;
SQ
Query Match 51.2%; Score 44; DB 1; Length 361;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 PWWPWR 11
DB 9 PQWPWR 15
RESULT 11
FUT3_PANTR STANDARD; PRT; 372 AA.
AC O19058;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Galactoside 3(4)-L-fucosyltransferase (EC 2.4.1.65) (Blood group Lewis
DE alpha-4-fucosyltransferase) (Lewis FT) (Fucosyltransferase 3) (FUCT-
DE III) (Alpha-3/4-fucosyltransferase).
GN FUT3.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS GLY-162 AND MET-304.
RX MEDLINE=98037800; PubMed=9368041;
RA Costache M., Apoll P.-A., Cailleteau A., Elmgren A., Larson G.,
RA Henry S., Blancher A., Iordachescu D., Oriol R., Mollicone R.;
RT "Evolution of fucosyltransferase genes in vertebrates.";
RL J. Biol. Chem. 272:29721-29728(1997).
CC -!- FUNCTION: May catalyze alpha-1,3 and alpha-1,4 glycosidic linkages
CC involved in the expression of sialyl Lewis x and Lewis x/SSEA-1
CC antigens. It may be involved in blood group Lewis determination
CC (By similarity).
CC -!- CATALYTIC ACTIVITY: GDP-beta-L-fucose + beta-D-galactosyl-(1->3)-
N-acetyl-D-glucosaminyl-R = GDP + beta-D-galactosyl-(1->3)-[alpha-
L-fucosyl-(1->4)]-N-acetyl-beta-D-glucosaminyl-R.
-!- PATHWAY: Glycosylation.
-!- SUBCELLULAR LOCATION: Type II membrane protein. Membrane-bound
form in trans cisternae of Golgi (By similarity).
-!- POLYMORPHISM: There are two alleles, A and B. Allele A has Arg-
162 and Val-304. Allele B has Gly-162 and Met-304.
-!- MISCELLANEOUS: Also acts on the corresponding 1,4-galactosyl
derivative, forming 1,3-L-fucosyl links.
-!- SIMILARITY: Belongs to the glycosyltransferase family 10.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL: Y14033; CAA74360.1;
DR InterPro; IPR001503; Glyco_transf_10.
DR Pfam; PF00852; Glyco_transf_10; 1.
KW Transfase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack; Polymorphism.
FT DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 15 34 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 35 372 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 162 162 R -> G (in allele B).
FT VARIANT 304 304 V -> M (in allele B).
FT SEQUENCE 372 AA; 43233 MW; 649CBF88CA7BD74C CRC64;
SQ
Query Match 51.2%; Score 44; DB 1; Length 372;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 PWWPWR 11
DB 9 PQWPWR 15
RESULT 12
MML6_MYCTU STANDARD; PRT; 397 AA.
AC Q10773;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative membrane protein mmpL6.
GN MML6 OR RV1557 OR M1608 OR MTCY48.08C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornaby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.

```



```

RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the mmpL family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z74020; CA98334.1; -.
DR EMBL; AF007027; AAK45875.1; -.
DR F1R; B70763; B70763.
DR TIGR; MT1608; -.
DR Tuberculist; Rv1557; -.
DR InterPro; IPR004869; MMP1.
DR Pfam; PF03176; MMP1; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 161 181 POTENTIAL.
FT TRANSMEM 190 210 POTENTIAL.
FT TRANSMEM 214 234 POTENTIAL.
FT TRANSMEM 242 262 POTENTIAL.
FT TRANSMEM 293 313 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
SQ SEQUENCE 397 AA; 42421 MW; 678DC86E244728F4 CRC64;

Query Match 51.2%; Score 44; DB 1; Length 397;
Best Local Similarity 75.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 RHPWFWPR 10
DB 351 RHPWFWPR 358

RESULT 13
YDM6-SCHPO
ID YDM6-SCHPO STANDARD; PRT; 535 AA.
AC O13912;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C23C11.06c in chromosome I.
GN SPAC23C11.06C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sources J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodson G.,
RA Holroyd S., Hornsby T., Howarth S., Rucke E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., Odell C.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,

```

```

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivy A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens J., Vanscreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochat M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Spakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z98559; CAB11159.1; -.
DR F1R; T38244; T38244.
DR GeneDB-SPombe; SPAC23C11.06c; -.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 82 102 POTENTIAL.
FT TRANSMEM 115 135 POTENTIAL.
FT TRANSMEM 143 163 POTENTIAL.
FT TRANSMEM 201 221 POTENTIAL.
FT TRANSMEM 346 366 POTENTIAL.
SQ SEQUENCE 535 AA; 60124 MW; A6AE149AAA292932 CRC64;

Query Match 51.2%; Score 44; DB 1; Length 535;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 1; Indels 4; Gaps 1;

QY 4 WPM-----WPMWR 11
DB 183 WPMSPSTWPMWRQ 194

RESULT 14
MML4-MYCTU
ID MML4-MYCTU STANDARD; PRT; 967 AA.
AC O53735;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative membrane protein mmpL4.
GN MML4 OR RV0450C OR MT0466 OR MV037.14C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:14:57 ; Search time 31.9947 Seconds
(without alignments)
118.710 Million cell updates/sec

Title: US-09-444-281-36

Perfect score: 86

Sequence: 1 ILRWPWPWRK 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_nhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	68.6	102	16	Q8P4Z9
2	59	68.6	105	16	Q8P4Z9 xanthomonas
3	57	66.3	723	12	Q9DUC4
4	53.5	62.2	137	10	Q84ST7
5	53.5	62.2	225	10	Q84ZK3
6	53	61.6	475	16	Q7U058
7	53	61.6	746	12	Q9JH31
8	53	61.6	1383	12	Q84712
9	53	61.6	1383	12	Q91AV1
10	53	61.6	1383	12	Q8B482
11	53	61.6	1386	12	Q8Q988
12	52	60.5	1245	3	Q9Y7V5
13	51	59.3	298	17	Q8ZU59
14	51	59.3	299	4	Q9Y4N1
15	51	59.3	351	16	Q8DJH5
16	51	59.3	504	2	P96143

ALIGNMENTS

RESULT 1

Q8P4Z9 ID Q8P4Z9 PRELIMINARY; PRT; 102 AA.
AC Q8P4Z9;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Inner membrane protein.
GN XCC3549.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=1204217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardoso J., Chambergo F., Ciprina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorfi H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira L.A.F., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.R.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing host specificities";
RL Nature 417:459-463(2002).
DR EMBL; AF012475; AAM42819.1; --
KW Complete proteome.
SQ SEQUENCE 102 AA; 11488 MW; 641654465C9571BF CRC64;

Q80VT9 mus musculus
Q9CZAI mus musculus
Q8PG47 xanthomonas
Q9AUN3 oryza sativ
Q7XFD1 oryza sativ
Q8DUC9 tt virus. o
Q7V8S5 prochloroco
Q9D80 tt virus. o
Q19573 caenorhabdi
Q988W4 rhizobium l
Q9D81 tt virus. o
Q91D04 tt virus. o
Q8PE93 xanthomonas
Q9X8C2 streptomyce
Q81GM8 sea may's su
Q8ZNS5 salmonella
Q8Z5Q0 salmonella
Q8G7C2 bifidobacte
Q84JN0 arabidopsis
O07504 bacillus me
Q8PHA9 xanthomonas
Q8V7E2 tt virus. o
Q9ENN3 oryza sativ
Q7U0Z3 mycobacteri
Q9HXC9 pseudomonas
Q7ULX9 rhodopirell
Q7V8E4 prochloroco
Q9U074 mycobacteri
Q8V711 tt virus. o

```

Query Match      68.6%; Score 59; DB 16; Length 102;
Best Local Similarity 87.5%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LRWPWPWP 9
DB 64 LRWPWPWP 71

RESULT 2
Q8PUP5 ID Q8PUP5 PRELIMINARY; PRT; 105 AA.
AC Q8PUP5;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Inner membrane protein.
GN XAC0590.

OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
CC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Canavan F., Cardoso J., Chamberg F., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.P., Franco M.C., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Seta J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Teai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL; AB011686; AM35479.1; -.
KW Complete proteome.
SQ SEQUENCE 105 AA; 11853 MW; 4DF5A59FBC5EF3C2 CRC64;

Query Match      68.6%; Score 59; DB 16; Length 105;
Best Local Similarity 87.5%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LRWPWPWP 9
DB 64 LRWPWPWP 71

RESULT 3
Q9DUC4 ID Q9DUC4 PRELIMINARY; PRT; 723 AA.
AC Q9DUC4;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE ORF1.
OS TT virus.
OC Viruses; ssDNA viruses; Circoviridae; Anellovirus.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mf-TTV9;
RA Okamoto H.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

```

```

[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=Mf-TTV9;
RX MEDLINE=20534983; PubMed=11080484;
RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,
RA Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;
RT "Species-specific TT viruses in humans and nonhuman primates and their
RT phylogenetic relatedness."
RL Virology 277:368-378(2000).
DR EMBL; AB041959; BAB19313.1; -.
DR GO; GO:0004185; P:serine carboxypeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001563; Peptidase_S10.
DR InterPro; IPR004219; TTVirus_Unk.
DR Pfam; PF02956; TT_ORF1; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
SQ SEQUENCE 723 AA; 85393 MW; 232D003098766344 CRC64;

Query Match      66.3%; Score 57; DB 12; Length 723;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PWPWPWR 11
DB 2 PWPWPWR 8

RESULT 4
Q84ST7 ID Q84ST7 PRELIMINARY; PRT; 137 AA.
AC Q84ST7;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Hypothetical protein OSUNBa0092N01.27.
GN OSUNBa0092N01.27.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Radresh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSUNBa0092N01 genomic sequence."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC120535; AAC73229.1; -.
KW Hypothetical protein.
SQ SEQUENCE 137 AA; 15795 MW; 67B62CFAD153CB99 CRC64;

Query Match      62.2%; Score 53.5; DB 10; Length 137;
Best Local Similarity 40.0%; Pred. No. 7.4;
Matches 8; Conservative 1; Mismatches 2; Indels 9; Gaps 1;

QY 1 ILRW-----PWPWPWR 11
DB 43 VTFWWRRLVRRPWPWPWR 62

RESULT 5
Q84ZR3 ID Q84ZR3 PRELIMINARY; PRT; 225 AA.

```

```

AC Q84ZK3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE OJ1372_D12.7 protein.
GN OJ1372_D12.7.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=ev. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
   clone:OJ1372_D12."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003827; BAC57651.1; -- 52096C5EA0083F77 CRC64;
SQ SEQUENCE 225 AA; 23825 MW; 52096C5EA0083F77 CRC64;

Query Match      62.2%; Score 53.5; DB 10; Length 225;
Best Local Similarity 61.5%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

Qy 3 RWPW---WPRRK 12
Db 111 RWCWAPWPRRR 123

RESULT 6
QYU058 PRELIMINARY; PRT; 475 AA.
AC QYU058;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Conserved hypothetical protein.
GN MB1413C.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmeyer K., Camus J.-C., Medina N., Mansoor H.,
   Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
   Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
   Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
   "The complete genome sequence of Mycobacterium bovis.";
   Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL: BX248338; CAD94274.1; --.
KW Complete proteome.
SQ SEQUENCE 475 AA; 51266 MW; 330A39CCFC6F90F4 CRC64;

Query Match      61.6%; Score 53; DB 16; Length 475;
Best Local Similarity 85.7%; Pred. No. 26;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LRWPWPP 8
Db 33 VRWPWPP 39

RESULT 7
QYU058 PRELIMINARY; PRT; 746 AA.
AC QYU058;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

```

```

DE ORF1.
OS TT virus.
OC Viruses; ssDNA viruses; Circoviridae; Anellovirus.
OX NCBI_TaxID=68887;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=TUN02;
RA Okamoto H.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=TUN02;
RX MEDLINE=20456801; PubMed=11003468;
RA Ukita M., Okamoto H., Nishizawa T., Tawara A., Takahashi M.,
   Iizuka H., Miyakawa Y., Mayumi M.;
   "The entire nucleotide sequences of two distinct TT virus (TTV)
   isolates (TUN01 and TUN02) remotely related to the original TTV
   isolates.";
   Arch. Virol. 145:1543-1559(2000).
RL Arch. Virol. 145:1543-1559(2000).
DR EMBL: AB028869; BAA94878.1; --.
DR InterPro: IPR004219; TTVirus_Unk.
DR Pfam: PF02956; TT_ORF1_1.
SQ SEQUENCE 746 AA; 88561 MW; E0B22953AE764E3E CRC64;

Query Match      61.6%; Score 53; DB 12; Length 746;
Best Local Similarity 54.5%; Pred. No. 39;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LRWPWPPWRRK 12
Db 1 MAWGWRWRRR 11

RESULT 8
Q84712 PRELIMINARY; PRT; 1383 AA.
AC Q84712;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Spike protein.
OS Porcine epidemic diarrhea virus (strain Br1/87) (PEDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=229033;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=94231173; PubMed=8176382;
RA Duarte M., Laude H.;
RT "Sequence of the spike protein of the porcine epidemic diarrhoea
   virus.";
   J. Gen. Virol. 75:1195-1200(1994).
RN [2]_
RP SEQUENCE FROM N.A.
RX MEDLINE=93389433; PubMed=8397280;
RA Bridgen A., Duarte M., Tobler K., Laude H., Ackermann M.;
RT "Sequence determination of the nucleocapsid protein gene of the
   porcine epidemic diarrhoea virus confirms that this virus is a
   coronavirus related to human coronavirus 229E and porcine
   transmissible gastroenteritis virus.";
   J. Gen. Virol. 74:1795-1804(1993).
RN [3]_
RP SEQUENCE FROM N.A.
RX MEDLINE=94120721; PubMed=8291230;
RA Duarte M., Tobler K., Bridgen A., Rasschaert D., Ackermann M.,
   Laude H.;
RT "Sequence analysis of the porcine epidemic diarrhoea virus genome
   between the nucleocapsid and spike protein genes reveals a polymo.";
   Virology 198:466-476(1994).
DR EMBL: Z25483; CAA80971.1; --.
DR InterPro: IPR002551; Corona_S1.
DR InterPro: IPR002552; Corona_S2.
DR Pfam: PF01600; Corona_S1_1.

```

DR Pfam: PF01601; Corona_S2; 1. Y --> N (IN REF. 1).
 FT CONFLICT 422 423
 SQ SEQUENCE 1383 AA; 151405 MW; 741C84D5DD3BDC4D CRC64;
 Query Match 61.6%; Score 53; DB 12; Length 1383;
 Best Local Similarity 62.5%; Pred. No. 67;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 2 LRWPWWPW 9
 Db 1321 IKWPWWVW 1328
 RESULT 9
 Q91AV1 PRELIMINARY; PRT; 1383 AA.
 AC Q91AV1
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Spike protein.
 OS Porcine epidemic diarrhea virus (strain CV777) (PEDV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=229032;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9339433; PubMed=8397280;
 RA Bridgen A., Duarte M., Tobler K., Laude H., Ackermann M.,
 RT "Sequence determination of the nucleocapsid protein gene of the
 RT porcine epidemic diarrhea virus confirms that this virus is a
 RT coronavirus related to human coronavirus 229E and porcine
 RT transmissible gastroenteritis virus.",
 RL J. Gen. Virol. 74:1795-1804(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94120721; PubMed=8291230;
 RA Duarte M., Tobler K., Bridgen A., Rasschaert D., Ackermann M.,
 RA Laude H.,
 RT "Sequence analysis of the porcine epidemic diarrhea virus genome
 RT between the nucleocapsid and spike protein genes reveals a polymorphic
 RT ORF.",
 RL Virology 198:466-476(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96112302; PubMed=8630538;
 RA Tobler K., Ackermann M.,
 RT "PEDV leader sequence and junction sites.",
 RL Adv. Exp. Med. Biol. 380:541-542(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98455678; PubMed=9782358;
 RA Bridgen A., Kocherhans R., Tobler K., Carvajal A., Ackermann M.,
 RT "Further analysis of the genome of porcine epidemic diarrhea virus.",
 RL Adv. Exp. Med. Biol. 440:781-786(1998).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Kocherhans R., Bridgen A., Ackermann M., Tobler K.,
 RT "The complete genome sequence of porcine epidemic diarrhea
 RT coronavirus.",
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF353511; AAK38656.1; -;
 DR InterPro: IPR002551; Corona_S1.
 DR InterPro: IPR002552; Corona_S2.
 DR Pfam: PF01600; Corona_S1; 1.
 DR Pfam: PF01601; Corona_S2; 1.
 SQ SEQUENCE 1383 AA; 151352 MW; 022B5E5E5435876D CRC64;
 Query Match 61.6%; Score 53; DB 12; Length 1383;
 Best Local Similarity 62.5%; Pred. No. 67;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 2 LRWPWWPW 9

Db 1321 IKWPWWVW 1328
 RESULT 10
 Q8B482 PRELIMINARY; PRT; 1383 AA.
 AC Q8B482
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Spike protein.
 OS Porcine epidemic diarrhea virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=28295;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Chirj99;
 RA Yeo S.-G., Krell P., Nagy E.,
 RT "Cloning and nucleotide sequence analysis of spike gene of porcine
 RT epidemic diarrhea virus detected in Korea.",
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AV167585; AN86621.1; -;
 DR InterPro: IPR002551; Corona_S1.
 DR InterPro: IPR002552; Corona_S2.
 DR Pfam: PF01600; Corona_S1; 1.
 DR Pfam: PF01601; Corona_S2; 1.
 SQ SEQUENCE 1383 AA; 151582 MW; B5BA4D7EE5371A54 CRC64;
 Query Match 61.6%; Score 53; DB 12; Length 1383;
 Best Local Similarity 62.5%; Pred. No. 67;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 2 LRWPWWPW 9
 Db 1321 IKWPWWVW 1328
 RESULT 11
 Q8QQ98 PRELIMINARY; PRT; 1386 AA.
 AC Q8QQ98
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Spike protein.
 GN SPK1
 OS Porcine epidemic diarrhea virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=28295;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kang T.-J., Lim Y.-Y., Jang Y.-S., Kwon T.-H., Kim D.-H., Yang M.-S.,
 RT "Spike protein gene of Korea Porcine Epidemic Diarrhea Virus.",
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF500215; AAM19716.1; -;
 DR InterPro: IPR002551; Corona_S1.
 DR InterPro: IPR002552; Corona_S2.
 DR Pfam: PF01600; Corona_S1; 1.
 DR Pfam: PF01601; Corona_S2; 1.
 SQ SEQUENCE 1386 AA; 151853 MW; 11F98BCB2AA0526B CRC64;
 Query Match 61.6%; Score 53; DB 12; Length 1386;
 Best Local Similarity 62.5%; Pred. No. 67;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 2 LRWPWWPW 9
 Db 1324 IKWPWWVW 1331

RESULT 12
Q9Y7V5 PRELIMINARY; PRT; 1245 AA.
ID Q9Y7V5
AC Q9Y7V5
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Conidiospore surface protein.
GN CMPI
OS Trichoderma harzianum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_TaxID=5544;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 32173;
RX MEDLINE=99343881; PubMed=10413618;
RA Puyesky M., Benhamou N., Ponce Noyola P., Bauw G., Ziv T.,
RA Van Montagu M., Herrera Estrella A., Horwitz B.A.;
RT "Developmental regulation of cmpl, a gene encoding a multidomain
RT Conidiospore surface protein of Trichoderma";
RL Fungal Genet. Biol. 27:88-99(1999).
DR EMBL; AJ13651; CAB40845.1; -;
DR HSP; P01180; IMPO.
DR GO; GO:0005923; C:tight junction; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR006188; Claudin reg.
DR InterPro; IPR001673; S.oid repeat.
DR ProDom; PD006869; S.oid repeat; 2.
DR PROSITE; PS01346; CLAUDIN; 1.
DR PROSITE; PS01346; CLAUDIN; 1.
SQ SEQUENCE 1245 AA; 135824 MW; 3249C749AFA0CDF8 CRC64;

Query Match 60.5%; Score 52; DB 3; Length 1245;
Best Local Similarity 60.0%; Pred. No. 82;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 3 RWPWWPWRK 12
Db 1185 RWQWSPWR 1194

RESULT 13
Q8ZU59 PRELIMINARY; PRT; 298 AA.
ID Q8ZU59
AC Q8ZU59;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dihydropterolate synthase.
GN PAE2937.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum".
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AB009902; AJ64549.1; -;
DR GO; GO:0004156; F:dihydropterolate synthase activity; IEA.
DR GO; GO:0009396; P:folic acid and derivative biosynthesis; IEA.
DR InterPro; IPR000489; Dhdropt_synth.
DR InterPro; IPR006390; DHPS.
DR Pfam; PF00809; Pterin_bind; 1.
DR TIGRfams; TIGR01496; DHPS; 1.
KW Complete proteome.
SQ SEQUENCE 298 AA; 32885 MW; 0A463F36739D3ED1 CRC64;

Query Match 59.3%; Score 51; DB 17; Length 299;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 3 RWPWWPWR 9
Db 209 QWPWWKW 215

RESULT 14
Q9Y4N1 PRELIMINARY; PRT; 299 AA.
ID Q9Y4N1
AC Q9Y4N1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP434C192.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Ansgorge W., Wirkner U., Meves H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AL096753; CAB46428.2; -;
DR FIR; T12505; T12505.
DR KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 299 AA; 34032 MW; 6B8DB60E6A88239A CRC64;

Query Match 59.3%; Score 51; DB 4; Length 299;
Best Local Similarity 85.7%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 5 PWPWWRR 11
Db 37 PWPWWR 43

RESULT 15
Q8DJH5 PRELIMINARY; PRT; 351 AA.
ID Q8DJH5
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Tlr1250 protein.
GN TLR1250.
OS Synchococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriuchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1".
RL DNA Res. 9:123-130(2002).
DR EMBL; AF005373; BAC08802.1; -;
DR InterPro; IPR001678; Sun_Nop1/Nop2.
DR Pfam; PF01189; Noll_Nop2_Sun; 1.
KW Complete proteome.
SQ SEQUENCE 351 AA; 38494 MW; 675046ADCBE7C835 CRC64;

Query Match 59.3%; Score 51; DB 16; Length 351;
Best Local Similarity 71.4%; Pred. No. 35;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LRWPWP 8

$\frac{1}{2}$

1 MKWPWP 7

Search completed: May 4, 2004, 15:22:10
Job time : 32.2281 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2004, 15:08:11 ; Search time 49.6053 Seconds
(without alignments)
74.047 Million cell updates/sec

Title: US-09-444-281-85
Perfect score: 99
Sequence: 1 ILPKWPWPWRR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	100.0	13	2 AAR30970	Broad spe
2	99	100.0	13	2 AAR78457	Aar78457 Indolicid
3	99	100.0	13	2 AAW66441	Aaw66441 Cationic
4	99	100.0	13	2 AAY24608	Aay24608 Indolicid
5	99	100.0	13	2 AAW87609	Aaw87609 Antimicro
6	99	100.0	13	3 AAY44666	Aay44666 Crosslink
7	99	100.0	13	3 AAY44324	Aay44324 Antimicro
8	99	100.0	13	3 AAY91740	Aay91740 Cationic
9	99	100.0	13	3 AAY91771	Aay91771 Amino aci
10	99	100.0	13	3 AAY57123	Aay57123 Amino aci
11	99	100.0	13	3 AAY57123	Aay57123 Naturally
12	99	100.0	13	3 AAY55056	Aay55056 Non-amida
13	99	100.0	13	3 AAY92794	Aay92794 Synthetic
14	99	100.0	13	4 AAB91842	Aab91842 Antimicro
15	99	100.0	13	4 ABP030382	Abp030382 Indolicid
16	99	100.0	13	4 ABP030383	Abp030383 Indolicid
17	99	100.0	13	5 ABB07699	Abb07699 Bovine ca
18	99	100.0	13	5 ABB81940	Abb81940 Peptide f
19	99	100.0	13	5 ABP59052	Abp59052 Peptide #
20	99	100.0	13	5 AAU90977	Aau90977 Transplan
21	99	100.0	13	5 ABB81249	Abb81249 Indolicid
22	99	100.0	13	5 ABB81261	Abb81261 Indolicid
23	99	100.0	13	5 AAC35561	Aac35561 L-indolic
24	99	100.0	13	6 ADA00505	Ada00505 Antimicro
25	99	100.0	13	6 ADA00504	Ada00504 Antimicro

26	99	100.0	13	6 ABU59617	Abu59617 Cationic
27	99	100.0	13	6 ABG76068	Abg76068 Human reg
28	99	100.0	13	6 AAE34433	Aae34433 Cow indol
29	99	100.0	13	6 ABR00800	Abr00800 Bioactive
30	99	100.0	13	6 ABR00815	Abr00815 Bioactive
31	99	100.0	13	6 ABR63788	Abr63788 Bovine ca
32	99	100.0	13	7 ADC73322	Adc73322 Bovine an
33	99	100.0	13	7 ADC98852	Adc98852 Synthetic
34	99	100.0	13	7 ADC98851	Adc98851 Synthetic
35	99	100.0	13	8 ADD35365	Add35365 Antimicro
36	99	100.0	14	3 AAY57118	Aay57118 Indolicid
37	99	100.0	14	3 AAY57143	Aay57143 Indolicid
38	99	100.0	15	2 AAW12879	Aaw12879 Antimicro
39	99	100.0	15	6 ABG73946	Abg73946 Cell wall
40	99	100.0	16	3 AAY57144	Aay57144 Indolicid
41	99	100.0	19	5 AAB47907	Aab47907 C-terminu
42	99	100.0	63	3 AAY44668	Aay44668 Poly-(Ind
43	99	100.0	63	3 AAY57142	Aay57142 Indolicid
44	99	100.0	144	5 ABB07706	Abb07706 Bovine pe
45	96	97.0	13	2 AAR78459	Aar78459 Indolicid

ALIGNMENTS

RESULT 1
AAR30970
ID AAR30970 standard; peptide; 13 AA.
XX
AC AAR30970;
XX
DT 25-MAR-2003 (revised)
DT 12-MAY-1993 (first entry)
XX
DE Broad spectrum antimicrobial indolicidin peptide.
XX
KW Tryptophan rich; microbial; microbistatic; inhibition.
XX
OS Bos taurus.
XX
PN WO9222308-A1.
XX
PD 23-DEC-1992.
XX
PF 10-JUN-1992; 92WO-US004920.
XX
PR 14-JUN-1991; 91US-00715271.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Sclated ME, Cullor JS;
XX
DR WPI; 1993-017896/02.
XX
PT Broad spectrum antimicrobial cpd. obtd. from bovine granulocytes -
PT comprises tryptophan rich peptide, pref. having low immunogenicity and
PT comprising proline rich peptide or carboxy terminal amide.
XX
PS Claim 2; Page 19; 29pp; English.
XX
CC The sequence is that of an indolicidin peptide which shows broad spectrum
CC antimicrobial activity and when administered to a host does not elicit an
CC immune response. It is effective against viruses, gram positive bacteria,
CC gram negative bacteria and fungi, including Staphylococcus aureus, Candida
CC Escherichia coli, Salmonella typhimurium, Listeria monocytogenes, Candida
CC albicans and Cryptococcus neoformans. It can be used as a therapeutic
CC agent, food preservative or disinfectant, e.g. to purify a water supply.
CC The peptide is pref. administered at an effective amt. of 0.5-500 ug/ml
CC final concentration. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 2; Length 13;

Best Local Similarity 100.0%; Pred. No. 1.4e-06; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKPWPWPWR 13
 |||||
 Db 1 ILPWKPWPWPWR 13

RESULT 2

AAR78457

ID AAR78457 standard; peptide; 13 AA.

XX AC AAR78457;
 XX DT 25-MAR-1996 (first entry)
 XX DE Indolicidin analog #4.

XX Indolicidin; microbicide; therapeutic agent; prophylactic;
 KW food preservative; disinfectant; medication; Gram positive bacteria;
 KW Gram negative bacteria; protozoa; yeast; fungi; viruses.

XX Synthetic.
 OS WO9522338-A1.
 XX PN 24-AUG-1995.
 XX PD 10-FEB-1995; 95WO-US001895.
 XX PF 16-FEB-1994; 94US-00197205.
 XX PR (REGC) UNIV CALIFORNIA.
 XX PA Selsted ME;
 XX PI WPI; 1995-302552/39.
 XX DR

XX Analogues of the tryptophan-rich peptide indolicidin - exhibiting broad
 PT spectrum antimicrobial activity and selectivity without undesirable side
 PT effects.

XX Claim 6; Page 27; 37pp; English.

XX The sequences represented by AAR78454-R78459 are indolicidin analogues.
 CC These analogues exhibit broad spectrum antimicrobial activity and have
 CC antimicrobial selectivity when compared to naturally occurring
 CC indolicidin. The antimicrobial activity of these analogues can be altered
 CC by incorporation of D-form, chemically altered or synthetic amino acids.
 CC These sequences can be incorporated into a pharmaceutical composition
 CC (e.g. as a liposome or non-liposome lipid complex carrier) for use in a
 CC microbicidal method. These sequences are active against Gram positive and
 CC negative bacteria, protozoa, yeast, fungi and viruses. They can be used
 CC as therapeutic agents, prophylactics, food preservatives, disinfectants
 CC or medications. These sequences are easily synthesised in an active and
 CC effective broad spectrum antimicrobial form with decreased undesirable
 CC side effects. Compared to naturally occurring indolicidin, these analogues
 CC show increased antimicrobial and decreased haemolytic activity. Peptide
 CC stability, and period of activity within the cell can be increased or
 CC decreased according to the incorporation of D- or L-form amino acids

XX SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 2; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKPWPWPWR 13
 |||||
 Db 1 ILPWKPWPWPWR 13

RESULT 3

AAW66441

ID AAW66441 standard; peptide; 13 AA.

XX AC AAW66441;
 XX DT 12-JAN-1999 (first entry)
 XX DE Cationic peptide indolicidin.

XX Indolicidin analogue; resistance; cationic peptide; antibiotic;
 KW bacterial infection; tolerance; antibacterial; microorganism; bacteria;
 KW fungus; parasite; virus.

XX OS Bos taurus.
 XX PN WO9840401-A2.
 XX PD 17-SEP-1998.
 XX PF 10-MAR-1998; 98WO-CA000190.
 XX PR 10-MAR-1997; 97US-0040649P.
 XX PR 20-AUG-1997; 97US-00915314.
 XX PR 26-SEP-1997; 97US-0060099P.
 XX PR 25-FEB-1998; 98US-00030619.
 XX PA (MICR-) MICROLOGIX BIOTECH INC.
 XX PI Fraser JR, West MHP, Menicol PJ;
 XX DR WPI; 1998-520800/44.

XX New indolicidin peptide analogues - useful for, e.g. enhancing activity
 PT of antibiotic or overcoming tolerance, acquired resistance or inherent
 PT resistance of microorganisms.

XX PS Disclosure; Page 10; 105pp; English.

XX AAW66393 to AAW66469 represent native cationic peptides from the present
 CC invention. The present invention describes compositions and methods for
 CC treating infection, especially bacterial infections. The compositions and
 CC methods use cationic peptides in combination with an antibiotic agent
 CC which are then administered to a patient to enhance the activity of the
 CC antibiotic agent, to overcome: (a) tolerance; (b) acquired resistance;
 CC and (c) inherent resistance. The combinations of antibiotics and cationic
 CC peptides can provide synergistic activity against a microorganism that is
 CC tolerant, inherently resistant, or has acquired resistance to an
 CC antibiotic agent. They can be used for killing e.g. bacteria, fungi,
 CC parasites and viruses

XX SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 2; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKPWPWPWR 13
 |||||
 Db 1 ILPWKPWPWPWR 13

RESULT 4

AAZ24608

ID AAZ24608 standard; peptide; 13 AA.

XX AC AAZ24608;
 XX DT 18-AUG-1999 (first entry)
 XX DE Indolicidin analogue #60.

XX Indolicidin; bacterial infection; photo-oxidised solubiliser;
 KW antimicrobial; antibiotic; antiarrhythmic; surface disinfectant; additive;


```

OS Bos sp.
XX Key Location/Qualifiers
FH Modified-site 6..9
FT FT /note= "Residues at positions 6 and 9 form a di-
FT tryptophan crosslink"
FT Modified-site 13
FT FT /note= "C-terminal amide"
FT FT
XX WO9965510-A1.
XX PN
XX XX
XX PD
XX PD
XX PF
XX PF
XX PF
XX PR
XX PR
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Selsted ME, Osapay K;
XX PI WPI; 2000-147133/13.
XX DR
XX XX
XX PT Crosslinked indolicidin analogs with antimicrobial activity against
XX bacteria, yeast, fungi, protozoa and viruses.
XX PT
XX PS Claim 3; Page 39; 53pp; English.
XX XX
XX CC The patent discloses crosslinked analogs of indolicidin (Indol 1-13)
XX CC which is a naturally occurring peptide isolated from bovine neutrophils
XX CC and has antimicrobial activity. The crosslinked indolicidin (X-
XX CC indolicidin) analogs are stable and have antimicrobial activity against
XX CC gram positive and negative bacteria (e.g. Staphylococcus aureus,
XX CC Escherichia coli; and Salmonella typhimurium), yeasts and fungi (e.g.
XX CC Candida albicans, Cryptococcus neoformans), protozoa (e.g. Giardia
XX CC species and Acanthamoeba species), and viruses (e.g. HIV-1). They can be
XX CC used for reducing or inhibiting the growth or survival of microorganisms
XX CC in an environment e.g. a food or food product, a solution, an inanimate
XX CC object comprising a surface, or a mammal. The present sequence is a
XX CC specifically claimed X-indolicidin analog, Indol 1-13(W6/9) which
XX CC contains a di-tryptophan crosslink
XX CC
XX SQ Sequence 13 AA;
XX
Query Match 100.0%; Score 99; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. NO. 1.4e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILPWKPWPWPWR 13
DB 1 ILPWKPWPWPWR 13
RESULT 7
AAY44324
ID AAY44324 standard; peptide; 13 AA.
XX AC
XX AC AAY44324;
XX XX
XX DT 29-FEB-2000 (first entry)
XX XX
XX DE Antimicrobial peptide, Indolicidin.
XX XX
XX KW purF gene; glutamine pyrophosphoribosyl pyrophosphate amidotransferase;
XX KW purF derivative; fusion partner; antimicrobial peptide; indolicidin;
XX KW mass production; cleavage site; hydroxylamine; CNBr; DNA construct; cow;
XX KW neutralise; toxicity; pharmaceutical industry; food industry.
XX KW
XX OS Bos taurus.
XX XX
XX PN WO9964611-A1.
XX XX
XX PD 16-DEC-1999.

```

```

XX 08-JUN-1999; 99WO-KR000282.
XX PF
XX 09-JUN-1998; 98KR-00022117.
XX PR
XX 14-MAY-1999; 99KR-00017920.
XX XX
XX (SAMY-) SAMYANG GENEX CORP.
XX PI
XX PI Kim JH, Kang MH, Lee J, Park SH, Lee JW, Hong SS, Lee H;
XX DR WPI; 2000-097542/08.
XX DR N-PSDB; AAZ44324.
XX XX
XX PT New DNA constructs useful for mass production of antimicrobial peptides
XX PT in microorganism hosts.
XX PS Claim 1; Fig 1; 67pp; English.
XX XX
XX CC The present amino acid sequence is an antimicrobial peptide, Indolicidin
XX CC derived from cow, Bos taurus. It is used along with a derivative of purF
XX CC gene sequence that functions as a fusion partner. A DNA construct that
XX CC comprises this antimicrobial peptide encoding sequence and the entire,
XX CC partial or derivative of purF gene, is used for mass production of the
XX CC antimicrobial peptide in microorganisms without killing the host cells.
XX CC Use of the purF gene derivative sequence, neutralises the toxicity of the
XX CC antimicrobial peptides against the host microorganism. The antimicrobial
XX CC peptides are useful commercially in the pharmaceutical and food
XX CC industries
XX XX
XX SQ Sequence 13 AA;
XX
Query Match 100.0%; Score 99; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. NO. 1.4e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILPWKPWPWPWR 13
DB 1 ILPWKPWPWPWR 13
RESULT 8
AAY91740
ID AAY91740 standard; peptide; 13 AA.
XX AC
XX AC AAY91740;
XX XX
XX DT 06-JUN-2000 (first entry)
XX XX
XX DE Cationic peptide Indolicidin amino acid sequence.
XX XX
XX KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
XX KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
XX KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
XX KW multidrug resistance.
XX KW
XX OS Unidentified.
XX XX
XX PN WO9965506-A2.
XX XX
XX PD 23-DEC-1999.
XX XX
XX PF 14-JUN-1999; 99WO-CA000552.
XX XX
XX PR 12-JUN-1998; 98US-0009541.
XX XX (MTCR-) MICROLOGIX BIOTECH INC.
XX PA
XX PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
XX PI WPI; 2000-223549/19.
XX DR
XX XX
XX PT Novel pharmaceutical composition containing optionally activated
XX PT polyoxyalkylene-modified cationic peptides, useful for treating tumors.

```

XX Disclosure; Page 11; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon

SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 3; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILPWKPWWPWR 13
 |||||
 DB 1 ILPWKPWWPWR 13

RESULT 9
 AAY91771
 ID AAY91771 standard; peptide; 13 AA.

AC AAY91771;

XX 06-JUN-2000 (first entry)

XX Amino acid sequence of cationic peptide MBI 10.

XX Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.

XX Synthetic.

XX WO9965506-A2.

XX 23-DEC-1999.

XX 14-JUN-1999; 99WO-CA000552.

XX 12-JUN-1998; 98US-00096541.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
 DR WPI; 2000-223549/19.

XX Novel pharmaceutical composition containing optionally activated
 PT polyoxalkylene-modified cationic peptides, useful for treating tumours.

XX Disclosure; Page 14; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon

XX Sequence 13 AA;

Query Match 100.0%; Score 99; DB 3; Length 13;

Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILPWKPWWPWR 13
 |||||
 DB 1 ILPWKPWWPWR 13

RESULT 10

AAAY91772
 ID AAY91772 standard; peptide; 13 AA.

XX AAY91772;

XX 06-JUN-2000 (first entry)

XX Amino acid sequence of cationic peptide MBI 10CN.

XX Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.

XX Synthetic.

XX WO9965506-A2.

XX 23-DEC-1999.

XX 14-JUN-1999; 99WO-CA000552.

XX 12-JUN-1998; 98US-00096541.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
 XX WPI; 2000-223549/19.

XX Novel pharmaceutical composition containing optionally activated
 PT polyoxalkylene-modified cationic peptides, useful for treating tumours.

XX Disclosure; Page 14; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon

XX Sequence 13 AA;

Query Match 100.0%; Score 99; DB 3; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILPWKPWWPWR 13
 |||||
 DB 1 ILPWKPWWPWR 13

RESULT 11

AAAY57123
 ID AAY57123 standard; peptide; 13 AA.

XX AAY57123;

XX 28-FEB-2000 (first entry)

DE Naturally occurring bovine indolicidin peptide Indol 1-13.
 XX Indolicidin analogue; antimicrobial activity; helminth; bacteria; virus;
 KW treatment; inhibit growth; micro-organism; contact lens solution;
 KW transgenic plant; surgical instrument; yeast; fungi; protozoa.
 XX
 CS Bos sp.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 13
 FT /note= "C-terminal amide"
 XX
 PN WO9558141-A1.
 XX
 PD 18-NOV-1999.
 XX
 XX
 XX 05-MAY-1999; 99WO-US009942.
 PF
 XX 12-MAY-1998; 98US-00076227.
 PR
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX Selsted ME;
 PI
 XX WPI; 2000-053028/04.
 DR
 XX New indolicidin analogs, active against bacteria, yeast, fungi, protozoa
 PT and virus, used for, e.g. treating infections.
 PT
 XX
 XX Example 1; Page 28; 62pp; English.
 PS
 XX This sequence is a naturally occurring indolicidin peptide. Peptides
 CC AA57109-Y57138 and AA57143-Y57148 are new indolicidin analogues, which
 CC have a homoserine residue and/or a truncated amino terminal region. The
 CC analogues have the following amino acid sequence: Xaa1-Xaa2-Xaa3-Xaa4-
 CC Xaa5-Xaa6-Pro-Xaa6-Xaa6-Xaa7-Xaa7-Xaa8 where: Xaa1 = Ile, Leu,
 CC Val, Ala, Gly or absent; Xaa2 = Ile, Leu, Val, Ala, Gly or absent; Xaa3 =
 CC Pro or absent; Xaa4 = Trp, Phe or absent; Xaa5 = Arg, Lys or absent; Xaa6
 CC = Trp or Phe; Xaa7 = Arg, Lys or absent; Xaa8 = homoserine (Hse), Met,
 CC Met-Xaa9-Met or absent, and Xaa9 = at least one amino acid; provided that
 CC if Xaa1 is present, Xaa8 = Hse, Met or Met-Xaa9-Met; and further provided
 CC that: if Xaa2 is absent, Xaa1 is absent; if Xaa3 is absent, Xaa1 and Xaa2
 CC are absent; if Xaa4 is absent, Xaa1, Xaa2 and Xaa3 are absent; and if
 CC Xaa5 is absent, Xaa1, Xaa2, Xaa3 and Xaa4 are absent. The indolicidin
 CC analogues can be used to create a fusion polypeptide consisting of the
 CC analogue linked to a peptide. The indolicidin analogues have
 CC antimicrobial activity against gram positive bacteria, gram negative
 CC bacteria, yeast, fungus, protozoa and viruses (e.g. HIV-1). They are also
 CC active against helminths. The analogues can be used for reducing or
 CC inhibiting growth or survival of a microorganism. They can be used for
 CC treating infections. They can also be included in a liquid such as water
 CC or an aqueous solution, e.g. contact lens solution. The analogues have
 CC potential uses in food products, and in objects such as the surface of an
 CC instrument used to prepare food or to perform surgery. Transgenic plants
 CC or animals useful in the food industry can be produced by introducing a
 CC nucleic acid molecule encoding an indolicidin analogue into the germline
 CC cells of such organisms
 XX
 SQ Sequence 13 AA;
 Query Match 100.0%; Score 99; DB 3; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILPWKWPWWPWR 13
 DB 1 ILPWKWPWWPWR 13
 RESULT 12
 AA55056
 ID AA55056 standard; peptide; 13 AA.
 XX

AC AA55056;
 XX 23-FEB-2000 (first entry)
 DT
 XX Non-amidated indolicidin peptide.
 DE
 XX Indolicidin; bactericin; sulphate-reducing bacteria; growth inhibitor;
 KW corrosion; degradation; metal; concrete; cement; dental implant; biofilm.
 XX
 OS Bacillus sp.
 XX
 PN WO9956553-A1.
 XX
 PD 11-NOV-1999.
 XX
 XX 03-MAY-1999; 99WO-US009675.
 PF
 XX 06-MAY-1998; 98US-00074037.
 PR
 XX 31-MAR-1999; 99US-00282277.
 PR
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX Wood TK, Jayaraman A, Barthman JC;
 PI WPI; 2000-052882/04.
 DR
 XX Inhibiting growth of sulfate-reducing bacteria using other bacteria,
 PT particularly for protection of metals and concrete.
 PT
 XX Example 4; Page 41; 84pp; English.
 PS
 XX This sequence represents the non-amidated indolicidin peptide. The
 CC invention relates to a method for inhibiting growth of sulphate-reducing
 CC bacteria (A) on a material (B) sensitive to corrosion or degradation, by
 CC applying to (B) a bacterium (C) that secretes a compound (1) able to
 CC inhibit growth of (A). The method is used to protect metal, concrete or
 CC cement against corrosion and degradation, but (B) can also be used to
 CC protect dental implants. (B) is present in an open or closed system (e.g.
 CC water cooling tower, liquid storage container, fuel tank, sewer or
 CC drainage system etc.) or part of a bridge or other structure. The method
 CC is more effective and less expensive than known methods for inhibiting
 CC (A), and reduces the amount of toxic chemicals released. Conventional
 CC biofilms of aerobic organisms tend to encourage growth of (A), and
 CC addition of (C) to the biofilm prevents this. A single application of (C)
 CC lasts for a long time, and (1) are produced exactly where they are
 CC required and inhibit (A) without significant impact on other organisms
 CC (this effect includes reducing resistance of (A) to conventional
 CC biocides, which may then be used in reduced amounts). If local damage to
 CC the biofilm occurs, the underlying material is still protected by
 CC diffusion of (1) from neighbouring areas
 XX
 SQ Sequence 13 AA;
 Query Match 100.0%; Score 99; DB 3; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILPWKWPWWPWR 13
 DB 1 ILPWKWPWWPWR 13
 RESULT 13
 AA52794
 ID AA52794 standard; peptide; 13 AA.
 XX
 XX AA52794;
 AC
 XX 29-AUG-2000 (first entry)
 DT
 XX Synthetic antimicrobial peptide, indolicidin.
 DE
 XX Magainin; antimicrobial; transgenic plant; protease degradation; Rev4;
 KW

KW indolicidin; protein production; reverse peptide.
 XX
 OS Bos taurus.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 13
 FT /note= "amidated"
 XX
 XX WO200026344-A1.
 XX
 XX 11-MAY-2000.
 PD
 XX
 PF 29-OCT-1999; 99WO-US025561.
 XX
 PR 30-OCT-1998; 98US-0106373P.
 PR 02-NOV-1998; 98US-0106537P.
 FT
 XX (INTE-) INTERLINK BIOTECHNOLOGIES LLC.
 PA (KENT) UNIV KENTUCKY RES FOUND.
 PA
 XX Everett NP, Li Q, Lawrence C, Davies MH;
 XX
 XX WPI; 2000-365597/31.
 DR
 XX Polypeptides for reducing proteolytic degradation of proteins
 FT administered to, or produced by a plant comprise indolicidin or its
 FT functional equivalents.
 PT
 XX Example 2; Page 15; 50pp; English.
 PS
 XX Indolicidin is a potent antimicrobial tridecapeptide, originally purified
 CC from cytoplasmic granules of bovine neutrophils. A reverse peptide, Rev4
 CC (AAY92796) of indolicidin was found to have increased stability against
 CC plant protease degradation. Expression of antimicrobial peptides in
 CC transgenic plants suffers a major limitation in that the foreign peptides
 CC are susceptible to rapid degradation by proteases. The invention concerns
 CC reducing the extent of protease degradation of a protein applied to, or
 CC produced by a plant by administering indolicidin, Rev4 or a functional
 CC equivalent to the plant. Transgenic plants expressing indolicidin and
 CC Rev4 are useful for production of the antimicrobial peptides.
 CC Compositions containing indolicidin and Rev4 are also useful for
 CC production of agronomically important proteins in plants
 XX
 XX Sequence 13 AA;
 SQ
 Query Match 100.0%; Score 99; DB 3; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILPWKWPWPWRR 13
 DB 1 ILPWKWPWPWRR 13
 |||||
 RESULT 14
 AAB91842
 ID AAB91842 standard; peptide; 13 AA.
 XX
 AC AAB91842;
 XX
 DT 22-JUN-2001 (first entry)
 XX
 XX Antimicrobial peptide SEQ ID NO:1018.
 XX
 KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
 KW blood component; modification; succinimidyl; maleimido group; amino;
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200069900-A2.

XX 23-NOV-2000.
 PD
 XX 17-MAY-2000; 2000WO-US013576.
 PF
 XX 17-MAY-1999; 99US-0134406P.
 PR 10-SEP-1999; 99US-0153406P.
 PR 15-OCT-1999; 99US-0159783P.
 XX
 PA (CONJ-) CONJUCHEM INC.
 XX
 XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
 PI
 XX WPI; 2001-112059/12.
 DR
 XX Modifying and attaching therapeutic peptides to albumin prevents
 PT peptidase degradation, useful for increasing length of in vivo activity.
 FT
 XX Disclosure; Page 528; 733pp; English.
 PS
 XX The present invention describes a modified therapeutic peptide (I)
 CC comprising a therapeutically active amino acid region (III) and a
 CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
 CC a less therapeutically active amino acid region (IV), which covalently
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
 CC factors and neurotransmitters, to protect them from peptidase activity in
 CC vivo for the treatment of various disorders. Endogenous therapeutic
 CC peptides are not suitable as drug candidates as they require frequent
 CC administration due to rapid degradation by peptidases in the body.
 CC Modifying and attaching therapeutic peptides to albumin prevents or
 CC reduces the action of peptidases to increase length of activity (half
 CC life) and specificity as bonding to large molecules decreases
 CC intracellular uptake and interference with physiological processes.
 CC AAS90829 to AAB92441 represent peptides which can be used in the
 CC exemplification of the present invention
 XX
 XX Sequence 13 AA;
 SQ
 Query Match 100.0%; Score 99; DB 4; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILPWKWPWPWRR 13
 DB 1 ILPWKWPWPWRR 13
 |||||
 RESULT 15
 AAB90382
 ID AAB90382 standard; peptide; 13 AA.
 XX
 AC AAB90382;
 XX
 DT 28-MAR-2003 (first entry)
 XX
 XX Indolicidin peptide SEQ ID NO 1.
 DE
 XX Indolicidin; ophthalmic; disinfection; contact lens; antimicrobial;
 KW Pseudomonas aeruginosa; Staphylococcus aureus; Serratia marcescens;
 KW Candida albicans; Fusarium solani.
 XX
 XX Unidentified.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 13
 FT /note= "C-terminal CONH2"
 FT
 XX WO200071175-A1.
 PN
 XX 30-NOV-2000.
 PD
 XX

```

PF 23-MAY-2000; 2000WO-US014608.
XX
PR 25-MAY-1999; 99US-00318195.
XX
PA (LARG-) LARGE SCALE BIOLOGY CORP.
PA (STRI-) SRI INT.
PA (REGC-) UNIV CALIFORNIA.
PA (WESL-) WESLEY-JESSEN CORP.
PA (TUSE-) TUSE D.
PA (MORT-) MORTMANS K.
PA (HOKA-) HOKAMA L A.
PA (SELS-) SELSTED M E.
PA (CHAP-) CHAPOY L L.
PA (QUIN-) QUINN M H.
XX
PI Tuse D, Mortelmans K, Hokama LA, Selsted ME, Chapoy LL, Quinn MH;
XX
DR WPI; 2001-080322/09.
XX
PT Ophthalmic composition for storing, cleaning, or disinfecting contact
PT lens, comprises indolicidin, and buffer having specified halide ion
PT concentration or Good's buffer.
XX
PS Claim 15; Page 68; 91pp; English.
XX
CC The invention relates to an ophthalmic composition (I) for storing,
CC cleaning, or disinfecting a contact lens, comprising an indolicidin
CC antimicrobial peptide and a buffer having a halide ion concentration less
CC than 0.85 weight%, based on the total weight of (I) or Good's buffer. (I)
CC is a multipurpose solution for care of a contact lens and is suitable for
CC contact lens disinfection, storage, cleaning, conditioning, rehydrating,
CC moistening and lubricating. (I) is useful for disinfecting the contact
CC lens or contact lens storage vessel such as contact lens vial, contact
CC lens case or a contact lens shipping package by contacting the lens or
CC vessel with a disinfecting solution comprising (I). (I) is useful for
CC packaging a contact lens involving sealing the lens in a container with
CC (I), where the contact lens is not autoclaved. (I) reduces the number of
CC pseudomonas aeruginosa, Staphylococcus aureus and Serratia marcescens
CC organisms by 3.0 logs or more within 4 hours and the number of Candida
CC albicans and Fusarium solani by 1.0 log or more within 18 hours. (I) is
CC self-preserving and requires no additional preservatives or
CC disinfectants. Since indolicidins are safe for topical application to the
CC eye, (I) enables immediate application of the contact lens to the eye
CC without the need for neutralisation, deactivation or washing any of the
CC components of (I). The present sequence is that of an indolicidin peptide
CC of the invention
XX
SQ Sequence 13 AA;
Query Match 100.0%; Score 99; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. NO. 1.4e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ILPWKWPWPWRR 13
Db 1 ILPWKWPWPWRR 13
Search completed: May 4, 2004, 15:19:40
Job time : 50.6053 secs

```